Critical Reviews in Biochemistry and Molecular Biology, 36(1):39–106 (2001)

Thermophilic Adaptation of Proteins

Reinhard Sterner[1] and Wolfgang Liebl[2]*

[1]Institut für Biochemie der Universität zu Köln, Otto-Fischer-Str. 12-14, D-50674 Köln, Germany, and [2]Institut für Mikrobiologie und Genetik der Georg-August-Universität Göttingen, Grisebachstr. 8, D-37077 Göttingen, Germany

Referee: Dr. Ruth Nussinov, Saic Frederick, Bldg. 469. 469, Room 151, Frederick, MD 21702–1201

* Corresponding author. Phone: +49-221-470 6432. FAX: +49-221-470 6731. E-mail: Reinhard.Sterner@Uni-Koeln.de

ABSTRACT: Hyperthermophilic organisms optimally grow close to the boiling point of water. As a consequence, their macromolecules must be much more thermostable than those from mesophilic species. Here, proteins from hyperthermophiles and mesophiles are compared with respect to their thermodynamic and kinetic stabilities. The known differences in amino acid sequences and three-dimensional structures between intrinsically thermostable and thermolabile proteins will be summarized, and the crucial role of electrostatic interactions for protein stability at high temperatures will be highlighted. Successful attempts to increase the thermostability of proteins, which were either based on rational design or on directed evolution, are presented. The relationship between high thermo-stability of enzymes from hyperthermophiles and their low catalytic activity at room temperature is discussed. Not all proteins from hyperthermophiles are thermostable enough to retain their structures and functions at the high physiological temperatures. It will be shown how this shortcoming can be surpassed by extrinsic factors such as large molecular chaperones and small compatible solutes. Finally, the potential of thermostable enzymes for biotechnology is discussed.

Table of Contents

1040-9238/01/$.50

I. HYPERTHERMOPHILIC LIFE

Cellular life depends on a highly complex metabolic network, which is based on the stabilities of macromolecules and of metabolic intermediates. Regarding the low thermal stability of metabolites such as NAD(P)/H or ATP (Leibrock et al., 1995), the upper temperature limit of life has been estimated to lie below 150°C (Stetter, 1999).

Hyperthermophiles are prokaryotes with an optimum growth temperature of 80°C or above. The maximum confirmed growth temperature is 113°C for the archaeon *Pyrolobus fumarii*, a facultatively aerobic chemolithotrophic hydrogen oxidizer. This organism was isolated by Stetter and colleagues from the hydrothermally heated wall of a deep sea vent ("black smoker") at the Mid-Atlantic Ridge (Blöchl et al., 1997). A significant proportion of *Pl. fumarii* cells taken from the exponential growth phase survive autoclaving for 1 h at 121°C, which demonstrates the extraordinary thermotolerance of this species. About 70 hyperthermophilic archaea have been described to date. In contrast, only very few hyperthermophilic bacteria have been cultivated, belonging to the orders *Thermotogales* and *Aquificiales*. The model organisms for the study of hyperthermophilic archaea and bacteria, that is, those for which most information is available to date, belong to the genera *Pyrococcus* and *Thermotoga*, respectively.

A. Habitats, Diversity, and Phylogeny

The main habitats for hyperthermophiles include marine and terrestrial hot environments. The most spectacular marine habitats are the hot "smoker" chimneys in the deep sea, where water of 200 to 350°C is injected from below the sea floor into the cold (2 to 4°C) sea water. The rock walls of these "black smokers" contain high concentrations of hyperthermophiles. Hyperthermophiles have also been isolated from active sea mounts and from heated sediments in shallow sea water. Terrestrial, natural environments include solfataric fields and hot springs in volcanic areas. Extremely thermophilic bacteria and archaea were also found in deep subsurface heated (~60 to 100°C) oil reservoirs in Alaska, the North Sea, and Siberia (Stetter et al., 1993; Slobodkin et al., 1999), but it cannot be excluded that these organisms were introduced by man on oil-prospecting activities. Finally, man-made habitats such as hot coal refuse heaps or power plants used for the exploitation of geothermal energy have been scoured successfully for the presence of extreme thermophiles and hyperthermophiles (Stetter, 1999). In addition to these exotic habitats, moderate or even extreme thermophiles can be isolated from quite common places in our daily surroundings, for example, thermophilic bacilli and clostridia from compost piles or *Thermus* species from hot water boilers.

Novel microbiological isolation techniques are now available for isolating and characterizing new extreme thermophiles. For example, the "optical tweezers" method (Huber et al., 1995) allows for the recovery of single prokaryotic cells for inoculation into sterile growth media. In contrast to classic enrichment techniques, the isolated cell is enabled to multiply without having to compete with other microorganisms in the sample. Cultivation-independent methods

based on the PCR-aided amplification and comparative analysis of 16S rRNA sequences have been applied to various terrestrial and marine hot water environments (Barns et al., 1994; Hugenholtz et al., 1998; Reysenbach et al., 2000; Yamamoto et al., 1998). These investigations have shown that the microbial diversity in extremely hot habitats can be remarkably high in some locations, while it may be limited in others. While certain hot environments mainly host *Archaea*, others have been shown to contain *Bacteria* as their major constituents (Reysenbach et al., 2000). For example, in a sediment sample from Obsidian Pool of Yellowstone National Park the ratio of *Bacteria* to *Archaea* + *Eucarya* has been estimated at 75:1, while a sediment collected from a different pool had a ratio of 1:1 (Hugenholtz et al., 1998).

Phylogenetic analyses based on 16S rRNA comparisons indicate that hyperthermophiles represent the deepest and shortest branches in the domains *Archaea* and *Bacteria* (Stetter, 1999), suggesting that they could have retained archaic characteristics in their biomolecules and metabolism. Recent molecular phylogenetic studies of high-temperature communities have identified further organisms that root deeply in the bacterial and archaeal domains (Hugenholtz et al., 1998) and have even led to the proposal of the new kingdom "Korarchaeota" within the *Archaea* (Barns et al., 1996). Members of the korarchaeotal phylotype may be widely spread in hot environments (Takai and Sako, 1999; Reysenbach et al., 2000). Their further characterization is hampered by the difficult task of growing them in the laboratory, although it has been possible to propagate korarchaeota in stable mixed laboratory cultures (Burggraf et al., 1997). It will be challenging to recognize their specific growth requirements such as oxygen tension, nutrient composition and concentration, removal of toxic products of metabolism and dependence on syntrophic partner organisms, and to obtain pure cultures of these microorganisms for the study of their constituents and metabolism.

B. Metabolism and Molecular Adaptations

Most hyperthermophiles cultivated to date are chemolithoautotrophic. These organisms can obtain their cell carbon from carbon dioxide and display respiratory metabolism, gaining energy from the oxidation of sulfur, sulfides, hydrogen, or ferrous iron. The most important electron acceptors are oxygen (aerobic/microaerophilic respiration), nitrate, ferric iron, sulfate, sulfur, or carbon dioxide (anaerobic respiration) (Stetter, 1999). Some of the chemolithotrophs, for example, certain *Thermoproteus*, *Pyrobaculum*, or *Pyrodictium* species, are facultative heterotrophs and can use organic nutrients. Other species are obligate heterotrophs that gain energy either via respiratory or fermentative types of metabolism (Stetter, 1996, 1999). Interestingly, novel metabolic pathways have been detected in some hyperthermophilic archaea. For example, glucose fermentation in *P. furiosus* proceeds via a modified Embden-Meyerhof-Parnas pathway involving a glucokinase and a phosphofructokinase, which are ADP-dependent as opposed to the classic ATP-dependent enzymes found elsewhere (Kengen et al., 1994, 1996; Tuninga et al., 1999). *Thermococcus zilligii* was shown recently to operate a novel glycolytic pathway involving the formation of formate from C-1 in glucose via cleavage of a six-carbon carboxylic acid (Xavier et al., 2000). Thus, it appears that hyperthermophiles are excellent sources for novel enzymes and metabolic pathways.

The molecular mechanisms allowing organisms to cope with extremely high temperatures are diverse, including specific ad-

aptations on the levels of (a) the structures and functions of macromolecules and cellular structures, (b) physiology and metabolism, and (c) regulation of gene expression and maintenance of DNA integrity. At least three points deserve special attention in this context, namely, membrane, DNA, and protein stability and function at extreme temperatures. With respect to the cell membranes, it has been suggested that the tetraether lipid monolayer membranes found in extremely thermophilic archaea are better suited for life at high temperature than the bacterial bilayer membranes composed of ester-type lipids: ether links are more resistant to temperature than ester links, and tetraether lipid membranes display an extremely low proton and ion permeability (van de Vossenberg et al., 1998). Various mechanisms that help to maintain the primary and secondary structure of DNA at the temperatures of hyperthermophilic growth have been discussed by Grogan (1998).

This review is focused on the thermostability of proteins. Under extreme conditions of temperature, pH, and pressure, amino acids can be damaged irreversibly by deamidation, β-elimination, hydrolysis, Maillard reactions, oxidation, and disulfide interchange (Daniel et al., 1996; Jaenicke, 1998). As a result, above the boiling point of water, the half-life of some amino acids is significantly shorter than the generation time of hyperthermophiles (Jaenicke and Böhm, 1998). Furthermore, the hydrolysis of peptide bonds sets theoretical limits to protein stability by directly affecting the integrity of the polypeptide chain at its most basic level, the linear succession of amino acid residues. Nevertheless, a number of enzymes from hyperthermophiles are stable and active at temperatures considerably exceeding the upper growth limit of the producing organism. For example, the α-amylase from *P. woesei* and the L-isoaspartyl methyltransferase from *T. maritima* have been reported to display resistance against thermal inactivation well above the maximum growth temperatures of the producing organisms, which are 105 and 90°C, respectively (Koch et al., 1991; Ichikawa and Clarke, 1998).

II. MEASURING THE STABILITY OF PROTEINS

Proteins from thermophilic and hyperthermophilic organisms are generally described in the literature to be very (thermo)stable. However, the term "stability" is used with quite different meanings. Therefore, it seems appropriate to begin this section with the various definitions of protein "stability".

A. Thermodynamic Stability

Thermodynamic stability is the difference in the free energies between the folded (F) and unfolded (U) states of a protein, which is termed here ΔG_{FU}. ΔG_{FU} can only be determined if the two-state model of unfolding applies:

$$F \underset{k_f}{\overset{k_u}{\rightleftharpoons}} U \qquad (1)$$

where F is the native, folded structure, and U is the reversibly unfolded state; k_u and k_f are the rate constants for unfolding and (re)folding, respectively. The two-state model implies that (a) the unfolding of F to U is reversible and (b) only F and U are populated, that is, significant amounts of folding intermediates are absent. From the equilibrium constant $K = k_u/k_f = [U]/[F]$, ΔG_{FU} can be calculated as follows:

$$\Delta G_{FU} = - RT\ln K \quad (2)$$

At the standard temperature of 25°C, ΔG_{FU} values are generally determined by the analysis of protein unfolding induced by denaturants such as urea or guanidinium chloride (Pace and Scholtz, 1997). The degree of unfolding that is induced by a given denaturant concentration [D] is measured by fluorescence and/or circular dichroism (CD) spectroscopy (Schmid, 1997). The normalized spectroscopic signals yield K and thus ΔG_{FU} (cf. Equation 2) as a function of [D]. Assuming a linear relationship between ΔG_{FU} and [D], both ΔG_{FU} in the absence of denaturant (ΔG_{FU} (H_2O)) and the m-value, which is a measure of the dependence of ΔG_{FU} on [D], can be determined (Pace and Scholtz, 1997):

$$\Delta G_{FU} = \Delta G_{FU} (H_2O) - m [D] \quad (3)$$

ΔG_{FU} values for mesophilic proteins are often in the range of 30 to 60 kJ/mol, which is equivalent to a few hydrogen bonds or hydrophobic interactions (Pfeil, 1998). At 25°C, ΔG_{FU} values of thermophilic proteins are not necessarily higher than those of mesophilic proteins. The temperature at which U and F are equally populated (K = 1 and $\Delta G_{FU} = 0$) is called the melting temperature T_m, which is often used as a measure of the thermal stability of a given protein.

In order to obtain ΔG_{FU} over a broad range of temperatures, the enthalpy change (ΔH) and heat capacity difference (ΔC_p) between U and F can be determined by analyzing the thermal unfolding process (Pace and Scholtz, 1997). To this end, differential scanning calorimetry (DSC) and spectroscopic techniques (mostly CD) are used, which provide complementary results. DSC yields the enthalpy change of unfolding, ΔH^{cal}, which is model-independent. In contrast, CD yields $\Delta H^{van't\ Hoff}$, which is based on the two-state model:

$$d(\ln K)/d(1/T) = - \Delta H^{van't\ Hoff}/R \quad (4)$$

Consequently, a value of $\Delta H^{cal}/\Delta H^{van't\ Hoff}$ close to unity proves the validity of the two-state model (Privalov, 1979).

ΔH increases with temperature, because U has a higher heat capacity than F:

$$d(\Delta H)/dT = \Delta C_p \quad (5)$$

For the determination of ΔG_{FU} as a function of temperature, ΔH must be determined only at a single temperature. This is generally the melting temperature T_m, at which ΔH is termed ΔH_m. ΔC_p is often determined from the slope of a plot of ΔH_m vs. T_m (cf. Equation 5). To this end, DSC measurements are performed at various pH values, which results in a variation of T_m and a concomitant variation of ΔH_m (Privalov and Khechinashvili, 1974).

With the knowledge of T_m, ΔH_m, and ΔC_p, ΔG_{FU} as a function of temperature is calculated with the Gibbs-Helmholtz equation:

$$\Delta G_{FU}(T) = \Delta H_m(1-T/T_m) - \Delta C_p[(T_m - T) + T \ln(T/ T_m)] \quad (6)$$

Taking into account the following relationship between ΔC_p, ΔH_m, and the entropy change after unfolding, ΔS:

$$\Delta S(T) = \Delta S_m - \Delta C_p(T_m - T)/ T_m = \Delta H/ T_m - \Delta C_p (T_m - T)/ T_m \quad (7)$$

Equation 6 can be rewritten:

$$\Delta G_{FU}(T) = \Delta H_m - T \Delta S_m - \Delta C_p[(T_m - T) + T \ln(T/ T_m)] \quad (8)$$

Figure 1 shows a schematic profile of the $\Delta G_{FU}(T)$ function. It is a skewed parabola and intersects the x-axis twice, indicating that unfolding occurs both at low and high temperatures. The T_m value of a pro-

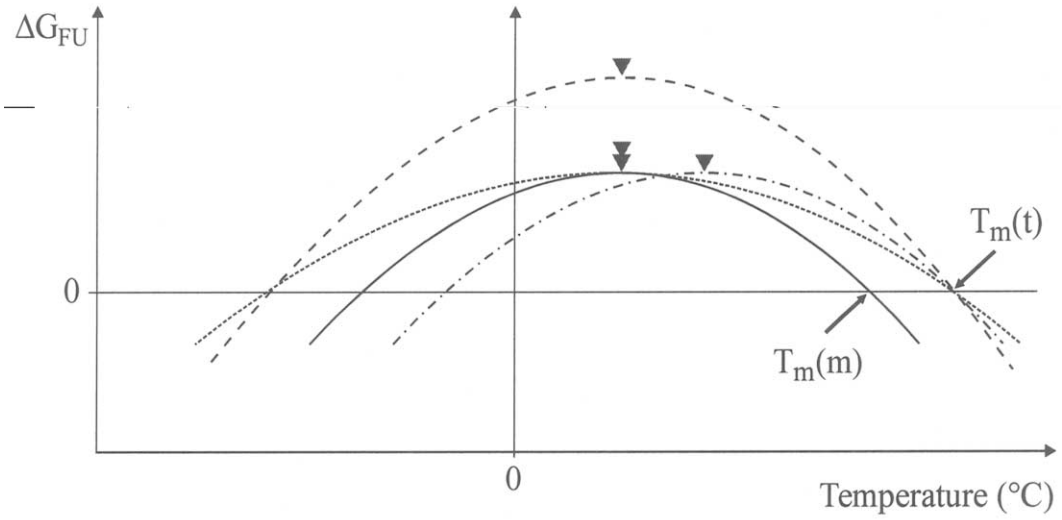


FIGURE 1. The T_m value of a protein can be increased by various strategies. Schematic representation of ΔG_{FU}, which is the difference in free energy of the unfolded and the folded state of a protein as a function of temperature. Compared with the ΔG_{FU} function of a hypothetical thermolabile protein (——) with $T_m(m)$, the ΔG_{FU} functions of hypothetical thermostable proteins can be up-shifted (– – – –), right-shifted (– · – · – ·), or flattened (-------), resulting in a higher T_m value ($T_m(t)$). ▼ indicates the temperature at which ΔG_{FU} reaches its maximum (T_{max}). Because the ΔG_{FU} functions are skewed parabolas, cold denaturation is also expected to occur, albeit normally far below 0°C.

tein can be increased by up-shifting, right-shifting or flattening of the $\Delta G_{FU}(T)$ function, or by any combination of the three.

The slope of the $\Delta G_{FU}(T)$ function is given by:

$$d(\Delta G)/dT = -\Delta S \qquad (9)$$

Therefore, at the temperature (T_{max}) of maximum thermodynamic stability: $\Delta S = 0$.

It follows from

$$\Delta G_{FU} = \Delta H - T\Delta S \qquad (10)$$

that at T_{max}, ΔG_{FU} is exclusively based on enthalpic contributions.

The curvature of the $\Delta G_{FU}(T)$ function is determined by ΔC_p, because

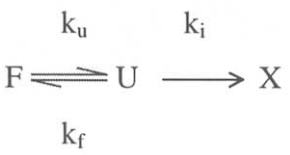
$$d^2(\Delta G)/dT^2 = -\Delta C_p/T \qquad (11)$$

In other words, the smaller ΔC_p, the flatter the curve and the higher the T_m (cf. Figure 1).

It has been shown that, to a first approximation, ΔC_p is a linear function of the number of amino acids of a protein (Myers et al., 1995), and that ΔC_p is correlated with the surface area of a protein that is hydrated after unfolding (= change in accessible surface area, ΔASA; Murphy and Freire, 1992). Because small proteins have a comparably smaller hydrophobic core than large proteins, the observed correlations of ΔC_p with the number of amino acids and with ΔASA highlight different views of the same physical phenomenon.

B. Kinetic Stability

The thermal unfolding of most mesophilic and thermophilic proteins is irreversible, mainly due to the aggregation of the unfolded state U, and therefore cannot be analyzed by equilibrium thermodynamics. A simple model for irreversible unfolding is



(12)

where X is the irreversibly unfolded state and k_i is the rate constant for the transition from U to X. In the case of irreversible unfolding, the measured rate of the overall process from F → X is termed $k_{obs}=(k_u \cdot k_i)/(k_f+k_i)$, and the half-life $t_{1/2}$ is given by $\ln 2/k_{obs}$. k_{obs} at a given temperature is used frequently as an operational measure of stability. Sometimes, also the half-inactivation temperature is used, that is the temperature at which half of the protein is unfolded after a given time of incubation. According to the model represented by Equation 12, the upper limit of k_{obs} is set by k_u, because only U (and not F) will undergo an irreversible transition. As Plaza del Pino et al. (2000) point out, this might be of special relevance for thermophilic proteins. At the high physiological temperatures of thermophiles irreversible processes, such as aggregation or covalent modification of thermally labile amino acids residues, will be fast and therefore $k_i >> k_f$ and $k_{obs} \sim k_u$. For this reason, a low k_u value is important to kinetically protect thermophilic proteins, especially *in vivo*. As we will see, in the few cases where k_u could be measured for a thermophilic protein, its value is much lower than that of the homologous mesophilic proteins.

C. Comparing the Stabilities of Mesophilic and Thermophilic Proteins

There is only a limited number of thermophilic proteins that undergo fully reversible unfolding and for which the $\Delta G_{FU}(T)$ function can be determined. Most of these proteins are small, monomeric, and consist of a single structural domain. In order to highlight their specific adaptation strategies, they will be compared here with their mesophilic homologues, as far as corresponding data are available. An increased thermodynamic stability of a given protein can be either due to a lower rate of unfolding and/or a higher rate of (re)folding. It is important to elaborate, under mechanistic and physiological aspects, which stabilization strategy was actually chosen by nature for a given protein. The few proteins for which the rates of unfolding and (re)folding could be measured are therefore discussed in somewhat more detail.

Sso7d from *Sulfolobus solfataricus* and Sac7d from *Sulfolobus acidocaldarius* are small DNA-binding proteins that consist of about 65 amino acid residues and show T_m values of 98 and 91°C at physiological pH values (Knapp et al., 1996; McCrary et al., 1996). Both Sso7d and Sac7d undergo reversible two-state unfolding. Their $\Delta G_{FU}(T)$ functions are shallow and the high T_m values are due to small ΔC_p values, which are, however, similar to the corresponding values of mesophilic proteins of comparable size. At physiological growth temperatures, both proteins are only marginally stable with a ΔG_{FU} of about 4 to 8 kJ/mol.

Ferredoxin from *T. maritima* is a small single-domain protein with 60 amino acid residues residues and a 4Fe-4S cluster. The cluster irreversibly decomposes in solution, rendering unfolding irreversible. Nevertheless, a thermodynamic analysis could be performed as all significant enthalpy changes are assumed to occur during the reversible unfolding of the polypeptide chain, which is before the irreversible loss of the cluster. The ΔG_{FU} (T) curve is both up- and right-shifted compared with thermolabile metalloproteins (cf. Figure 1), resulting in a T_m value of 125°C (Pfeil et al., 1997).

RNAsesH from *Thermus thermophilus* and *Escherichia coli* consist of about 160 amino acid residues with a sequence iden-

tity of 52% and an rms deviation for the Cα atoms in secondary structural elements of only 0.95 Å. This pair of closely related mesophilic and thermophilic enzymes is well suited for a comparative thermodynamic investigation (Hollien and Marqusee, 1999). To avoid unwanted cysteine chemistry, all cysteine residues in both proteins were replaced by serine or alanine. The corresponding variants were reversibly unfolded by GdmCl, yielding ΔG_{FU} values at various temperatures. Fitting of these data points to the Gibbs-Helmholtz equation (Equation 6) yielded ΔC_p, ΔH, and T_m, and allowed to generate the entire $\Delta G_{FU}(T)$ function. Interestingly, although both enzymes are of similar size and, according to their crystal structure, should undergo similar changes in accessible surface area after unfolding (ΔASA), ΔC_p is significantly lower for the thermophilic RNAseH (7.5 kJ/mol/K) than for its mesophilic counterpart (11.3 kJ/mol/K). This results in a flatter $\Delta G_{FU}(T)$ curve for *T. thermophilus* RNAse, which is additionally up-shifted. Consequently, the T_m value of the thermophilic enzyme is 20°C higher than that of from the mesophile (86 vs. 66°C). Remarkably, at the respective physiological temperatures of *T. thermophilus* and *E. coli*, the two enzymes show similar ΔG_{FU} values.

A particularily interesting and well-studied monomeric protein is rubredoxin from *Pyrococcus furiosus* (pRdx), which consists of 53 amino acid residues and an Fe^{2+} ion that is complexed by four cysteine residues. As for *T. maritima* ferredoxin referred to above, thermal unfolding of pRdx is irreversible, but its thermodynamic stability can be measured by hydrogen exchange in D_2O under native conditions (Hiller et al., 1997). The structural basis for these measurements is that hydrogen bonds are continuously formed and broken, even in the native state of a protein. The amide protons of hydrogen bonds in the so-called *closed* conformation are protected from exchange with deuterons in the bulk solvent. However, hydrogen bonds are broken in the *open* conformation, allowing the irreversible exchange of amide protons with solvent deuterons according to the following scheme:

$$\text{closed} \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} \text{open} \xrightarrow{k_{ch}} \text{exchanged} \qquad (13)$$

k_{op}, k_{cl}, and k_{ch} are the rate constants for the opening, the closing and the exchange reaction. Under so-called EX_2 conditions, where the closing reaction is much faster than the exchange reaction ($k_{cl} >> k_{ch}$), there is a preequilibrium between the closed and the open conformation that is described by the equilibrium constant K_{op}=[open]/[closed]=k_{op}/k_{cl}. The observed overall exchange rate is given by $k_{ex}=K_{op}k_{ch}$. From the so-called amide protection factor $1/K_{op}=k_{ch}/k_{ex}$ the free energy of the hydrogen bond opening reaction (ΔG_{HX}) can be calculated according to

$$\Delta G_{HX} = -RT\ln K_{op} \qquad (14)$$

Hiller et al. (1997) replaced the paramagnetic Fe of native pRdx by diamagnetic Zn, which is favorable for following the exchange of assigned hydrogen atoms by NMR. Under the assumption that the slowest exchanging amide protons, namely, those of the metal-complexing cysteines, exchange by global unfolding of pRdx, the minimum global stability ΔG_{FU} of the protein is equal to ΔG_{HX} of these hydrogens. The corresponding $\Delta G_{FU}(T)$ function yielded a thermodynamic stability of 63 kJ/mol at 100°C and an extrapolated T_m value of about 185°C, which makes pRdx the most thermostable protein investigated to date. In comparison to exceptionally thermostable proteins of nearly identical size from mesophiles, the

stability curve is considerably up-and right-shifted (cf. Figure 1). When pRdx is thermally unfolded at pH 7, an apparent (non-equilibrium) T_m value of about 110°C is measured (Cavagnero et al., 1995). This discrepancy between the two T_m values means that irreversible unfolding occurs at temperatures where U is not significantly populated, suggesting that k_u is rate limiting for the overall process from N to X (cf. Equation 12). In accordance with this hypothesis of kinetic protection, at pH 7 and 100°C, the overall rate constant constant of irreversible unfolding of pRdx is only about $2 \cdot 10^{-6} s^{-1}$ (Cavagnero et al., 1998).

The cold-shock proteins (Csp) from the mesophile *Bacillus subtilis* (*Bs*-Csp), the thermophile *Bacillus caldolyticus* (*Bc*-Csp), and the hyperthermophile *T. maritima* (*Tm*-Csp) are monomeric, all-β proteins that bind to single-stranded nucleic acids. They contain about 67 amino acid residues, 42 of which are identical in all three proteins. Both thermal- and guanidinium chloride induced unfolding are reversible, and the observed T_m values in 0.1 *M* salt at neutral pH are 50°C for *Bs*-Csp, 72°C for *Bc*-Csp and 85°C for *Tm*-Csp (Perl et al., 1998). The rate constants k_f of folding at 25°C are similar for the three proteins and range between 565 s^{-1} and 1380 s^{-1}. In contrast, unfolding kinetics are very different: k_u = 10 s^{-1} for *Bs*-CspB, 0.6 s^{-1} for *Bc*-Csp, and 0.018 s^{-1} for *Tm*-Csp. These results support the hypothesis that thermophilic and hyperthermophilic proteins are kinetically stabilized compared with their mesophilic counterparts.

There are at least two well-documented cases in the literature, where the thermodynamic stability of homologous *dimeric* proteins from organisms of quite different growth temperatures were compared. These are the archaeal histones (Li et al., 1998) and dihydrofolate reductases from *T. maritima* and *E. coli* (Dams and Jaenicke, 1999).

The histones from the mesophile *Methanobacterium formicium* (optimum growth at 43°C; HFoB), and from the hyperthermophiles *Methanothermus fervidus* (optimum growth at 83°C; HMfA and HMfB**)** and *Pyrococcus furiosus* (optimum growth at 100°C; HPyA1) are homodimers of subunits with about 68 amino acids and about 82% sequence identity. They unfold according to the following scheme:

$$F_2 \; (\rightleftharpoons 2F) \rightleftharpoons 2U \qquad (15)$$

that is, structured monomers F do not accumulate (Li et al., 1998). The stability of these histones was compared in terms of temperature, pH, and salt. Because the stability of reversibly unfolding oligomeric proteins is concentration dependent, the T_m and T_{max} values were compared at an extrapolated subunit concentration of 1 *M*. At pH 5 and 0.2 *M* salt, the following T_m/T_{max} values were determined: 75°C/32°C (HFoB), 104°C/35°C (HMfA), 113°C/40°C (HMfB), 114°C/44°C (HPyA1). The ΔC_p values are similar in the four proteins and the higher T_m values of the hyperthermophilic proteins are achieved by an up- and right-shifting of the $\Delta G_{FU}(T)$ function (Figure 2). Although the difference between the maximum thermodynamic stabilities of HPyA1 and HfoB amounts to 42 kJ/mol, all four histones are comparably stable at the corresponding optimum growth temperatures in accordance with the concept of corresponding states (Jaenicke, 1991). At low salt, HFoB, HMfA, and HMfB are far more labile. Along these lines, the three host organisms have high intracellular salt concentrations.

Dihydrofolate reductase is a homodimer in *T. maritima* (tDHFR), but a monomer in most mesophilic organisms. As observed for the archaeal histones,

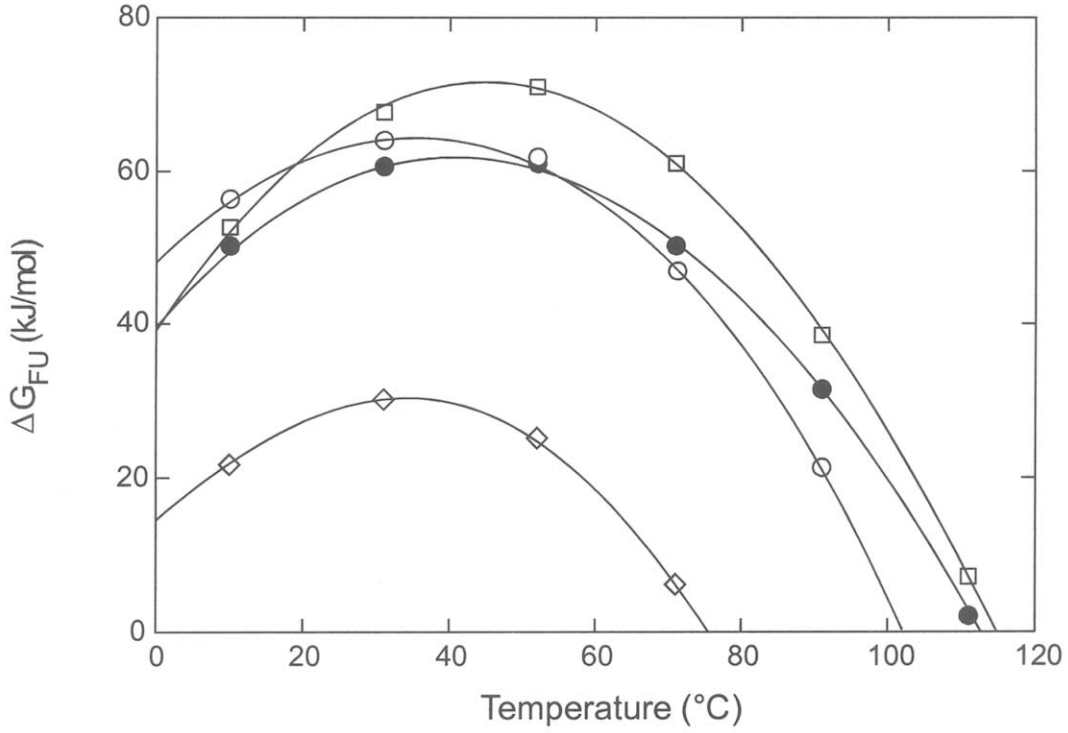


FIGURE 2. Hyperthermophilic histone proteins have higher ΔG_{FU} and T_m values than their mesophilic homologues. ΔG_{FU} functions of hyperthermophilic HPyA1 (□), HMfA (○), HMfB (•), and mesophilic HFoB (◇), at pH 5 in 0.2 *M* KCl. (Reprinted with permission from Li, W., Grayling, R.A., Sandman, K., Edmondson, S., Shriver, J.W., and Reeve, J.N. *Biochemistry* **37:** 10563–10572. Copyright (1998) American Chemical Society.)

tDHFR unfolds according to Equation 15 (Dams and Jaenicke, 1999), and T_{max} at an extrapolated subunit concentration of 1 *M* is 35°C. The $\Delta G_{FU}(T)$ function is not flattened, but up- and right-shifted compared with that of mesophilic homologues, as observed for pRdx and the hyperthermophilic histones. The unfolding of tDHFR is extremely slow, at 15°C $k_u = 4.6 \cdot 10^{-12}$ s^{-1}, which is about 10^8 times less than observed with the monomeric DHFR from *E. coli* (eDHFR) under identical conditions. Even at 55°C, k_u of tDHFR is 10^5 times smaller than k_u of eDHFR at 15°C. In contrast, the folding rates k_f of tDHFR and eDHFR are in the same order of magnitude, again supporting the concept of kinetic stabilization of thermophilic and hyperthermophilic proteins.

D. Summary

The $\Delta G_{FU}(T)$ function of a protein determines its T_m value, which is the temperature at which half of the protein molecules are native and the other half are unfolded. Nature has used all possible strategies to increase the T_m value of thermophilic proteins, namely, the up-shifting of the $\Delta G_{FU}(T)$ function to higher ΔG_{FU} values, its right-shifting to higher temperatures, and its flattening due to a decreased ΔC_p value of unfolding. In some thermophilic proteins one of these strategies dominates, in others a combination of two of them is observed. There are only few proteins, for which both the rates of unfolding and (re)folding could be measured. In all cases, the higher thermodynamic stability of the thermophilic

proteins appears to be due to a decreased rate of unfolding.

III. ELUCIDATING THE STRUCTURAL BASIS OF PROTEIN THERMOSTABILITY

It was shown in the previous section that proteins from (hyper-) thermophiles in general are intrinsically far more stable than proteins from mesophiles. There is enormous interest to reveal the structural basis of these stability differences, not only to understand mechanisms of high-temperature adaptation but also to be able to design proteins with a desired combination of stability and function (Hough and Danson, 1999).

Numerous studies have identified a number of potential interactions that nature may use to render proteins extremely thermostable (Jaenicke and Böhm, 1998). Among them are (1) an increase in the number of hydrogen bonds, (2) additional or improved electrostatic interactions caused by salt bridges or networks thereof, (3) optimized hydrophobic interactions, (4) increased compactness or packing densities, (5) increased polar compared with non-polar surface areas, (6) increases in α-helical content and α-helix stability, (7) the (improved) binding of metal ions, (8) improved fixation of the polypeptide chain termini to the protein core, (9) replacement of residues with energetically unfavourable conformations by glycine, (10) truncation of solvent-exposed loops, (11) a higher number of prolines and β-branched amino acids in loops, (12) association to oligomers, (13) reduction of the content of the thermally labile amino acids asparagine, glutamine, cysteine and methionine. This list shows that stabilizing features can occur at all structural levels, from the amino acid sequence to the quaternary structure of proteins. It is evident that some of these features are mutually interdependent; for example, a dimeric protein has a larger change in solvent-accessible surface area after unfolding (ΔASA per subunit), compared with a monomeric one.

The potentially stabilizing features listed above were mainly identified by comparing the X-ray structures of homologous pairs of mesophilic and (hyper)thermophilic proteins. However, the many neutral mutations that accumulate during divergent evolution make it difficult to identify the crucial amino acid differences by mere inspection of structures. Furthermore, the presence or absence of few stabilizing increments might account for significant stability differences (Jaenicke et al., 1996). Another difficulty arises from the temperature-dependence of stabilizing interactions. It was suggested that hydrophobic interactions, which are entropic at room temperature but become enthalpic at higher temperatures, reach their maximum stabilizing effect at 75°C (Makhatadze and Privalov, 1995). Recent theoretical studies suggest that the stabilizing effect of electrostatic interactions increases with increasing temperature (Elcock, 1998; de Bakker et al., 1999; Xiao and Honig, 1999; cf. Section D). Data on the temperature dependence of other stabilizing interactions are sparse.

For these reasons, it is not yet possible to derive rules that govern high protein thermostability. As outlined above, the identification of stabilizing interactions is a search for small differences against a huge background. Consequently, only large-scale comparisons of amino acid sequences and three-dimensional structures will provide statistically significant differences. The growing amount of whole genome sequences from mesophilic and (hyper)thermophilic organisms and the enormous speed with which new X-ray structure become available, now allow systematic comparisons between proteins from mesophiles and

(hyper)thermophiles that promise a more general insight into the problem.

A. Comprehensive Comparisons of Amino Acid Sequences and Three-Dimensional Structures

1. Comparison of Amino Acid Composition and Sequences between Entire Genomes

Recently, in several studies the amino acid compositions of mesophilic and hyperthermophilic proteins were compared on the level of whole genomes. Haney et al. (1999) aligned the sequences of 115 proteins from the hyperthermophile *Methanococcus janaschii*, which has an optimum growth temperature of 85°C, with homologous proteins from several mesophilic *Methanococcus* species, which have optimum growth temperatures of about 35°C. The advantages of this approach are that (1) multiple, diverse proteins are compared to diagnose common features; (2) the high sequence identity of about 70% between the compared proteins minimizes the noise that is due to phylogenetic differences; (3) the very different growth temperatures emphasize the essential properties; (4) the similar G + C content of the compared genomes minimizes potential bias due to different amino acid usage. The results are summarized in Table 1. Most significantly, there is a decrease in hyperthermophilic proteins in the content of the uncharged polar residues Gln, Asn, Thr and Ser, which results in a net loss of nine of these residues in a typical (300 residues) polypeptide chain. Gln and Asn residues are prone to deamidation, which can be catalyzed by Ser and Thr residues of the protein (Wright, 1991). In contrast, hyperthermophilic proteins show an increase in the content of charged amino acid residues, especially Glu, Arg and Lys, which results in a net gain of 6.5 charged residues in a typical protein. The described trends apply to about 88% of the compared proteins and were essentially confirmed by Cambillau and Claverie (2000), who compared the amino acid compositions of proteins from 22 mesophiles and seven hyperthermophiles. Their analysis of a large number of three-dimensional structures suggests that the extra charged residues are located in solvent-accessible regions at the surfaces of the hyperthermophilic proteins.

Chakravarty and Varadarajan (2000) note that the average thermophilic protein contains 268 ± 38 amino acid residues and is significantly smaller than the average mesophilic protein, which contains 310 ± 16 residues. This could be a significant finding, because small proteins show a lower heat capacity change after unfolding, ΔC_p, compared with larger ones (cf. Section II.A). Low ΔC_p values result in a flattening of the $\Delta G_{FU}(T)$ function (Equation 11) and, as a consequence, an increase of T_m. Thompson and Eisenberg (1999) also find that thermophilic proteins are shorter than their mesophilic homologues, due to shorter solvent-exposed loops. General thermodynamic considerations suggest that shortening of solvent-exposed loops would lower the entropy gain after unfolding and increase the T_m value. The main point is that a decrease in the conformational entropy of the *native* state by loop shortening would be small and restricted to the loop itself. In contrast, loop shortening would have a multiplicative effect on the number of conformations for the entire chain in the *unfolded* state.

2. Systematic Comparisons of Structures

In a number of comprehensive studies the three-dimensional structures of thermo-

TABLE 1
Change in Amino Acid Composition Going from Proteins of Mesophiles to Proteins of Thermophiles

Amino acid	Gains	Losses	Ratio	Net change	Change,%
Ile	842	658	1.28	184	9.5
Glu	739	562	1.31	177	9.1
Arg	383	214	1.79	169	16.5
Lys	789	620	1.27	169	8.3
Pro	167	96	1.74	71	7.0
Tyr	224	177	1.27	47	5.8
Ala	504	458	1.10	46	2.8
Trp	23	11	2.09	12	8.3
Leu	560	548	1.02	12	0.6
Cys	72	69	1.04	3	0.9
Phe	200	202	0.99	-2	-0.3
Asp	429	432	0.99	-3	-0.2
Val	666	670	0.99	-4	-0.2
His	80	92	0.87	-12	-2.8
Gly	201	264	0.76	-63	-3.4
Met	174	248	0.70	-74	-11.3
Gln	158	234	0.68	-76	-13.1
Thr	336	431	0.78	-95	-8.4
Asn	313	481	0.65	-168	-15.9
Ser	271	664	0.41	-393	-31.7

philic and mesophilic proteins were compared with respect to different criteria. Szilagyi and Závodszky (2000) generated a nonredundant data set containing all available high-quality X-ray structures of monomers and subunits of thermophilic proteins and their mesophilic counterparts. Subunit-subunit interactions in oligomeric proteins were ignored. The resulting data set contained 29 thermophilic and 64 mesophilic subunit structures that were compared with respect to hydrogen bonds, secondary structures, polarity of surfaces, cavities, and ion pairs. The proteins were catagorized in 25 different structural families, each of which contained at least one thermophilic and one mesophilic protein. A distinction was made between proteins from thermophiles with optimum growth temperatures between 45 and 80°C ($S_{45\text{-}80}$; 24 structures) and those with optimum growth temperatures of about 100°C (S_{100}; five structures). A given property of a particular thermophilic protein was compared with the average property of its mesophilic homologues. The most important results are as follows. There is no sig-

nificant difference between mesophilic and thermophilic proteins with respect to the (normalized) number of hydrogen bonds and the (normalized) number of unsatisfied hydrogen-bond donors plus acceptors. Only a weak positive correlation between growth temperature and secondary structure content was observed, and the ratio of polar to apolar surface areas did also not differ significantly between mesophilic and thermophilic proteins. However, the total surface area of cavities and especially their (normalized) number is significantly decreased in S_{100}, but unchanged in $S_{45\text{-}80}$. The most prominent difference between mesophilic and thermophilic proteins was detected with respect to ion pairs. The authors divided ion pairs into different categories according to the maximum distances between the positively charged nitrogen and the negatively charged oxygen atom (N-O distance) of the involved basic and acidic amino acid side chains. The N-O distance limits were set to 4.0, 6.0, and 8.0 Å, corresponding to a "weakening" of the electrostatic attractions in this order. Potential ion pairs involving histidine residues were not considered, because it is difficult to decide whether a given histidine side chain is charged in a protein. A significant increase in the (normalized) number of ion pairs is observed in thermophilic proteins; in $S_{45\text{-}80}$ weaker ion pairs dominate, whereas in S_{100} also stronger ion pairs are important (Table 2). From the correlation of the number of ion pairs with the optimum growth temperature, 4 additional strong and 14 additional weaker ion pairs are expected in a 300-residue protein from an organism growing at 80°C compared with that from an organism growing at 30°C. When the protein families are analyzed individually, it seems that most stabilizing features are used only in some families, that is, each protein family uses a specific combination of stabilizing interactions. It is not clear, why $S_{45\text{-}80}$ and S_{100} proteins use somewhat different strategies. This could be due to the temperature dependence of the individual interactions and/or due to phylogeny, as all S_{100} are from archaea.

In a similar investigation, Kumar et al. (2000a) created a nonredundant data set of 18 different protein families, each of which contained homologous mesophilic and thermophilic representatives. In contrast to the study of Szilagyi and Závodszky (2000), the comparison also included subunit-subunit interactions. The proteins were compared according to the following criteria: compactness or packing density (ratio of solvent accessible area of the protein to surface area of a sphere with the same volume as the protein; the smaller the value, the more compact or the better packed is a protein); hydrophobicity (fraction of the buried *non-polar* area out of the total *non-polar* area); polar and non-polar surface areas (fraction of the polar *exposed* area out of the total *exposed* area; and fraction of the polar *buried* area out of the total *buried* area); protein size (as reflected in different oligomeric states or insertion/deletion of residues); Pro residues in loops; helical content; helix stability (as inferred from the fraction of helix-favouring and helix-disfavoring residues); main chain-main chain, main chain-side chain and side chain-side chain hydrogen bonds, ion pairs with an N-O distance cut-off of 4 Å; amino acid composition. From each family the structurally most similar thermophile-mesophile pair with the highest resolution was compared. For most of the applied criteria there were no consistent trends among the families. The compactness, hydrophobicities, and polar and non-polar surface areas are not different, both for monomeric and oligomeric mesophilic-thermophilic protein pairs. Also, there was no significant difference detected with respect to the protein size and the number of Pro residues in loops. However, thermophilic proteins contain a higher

TABLE 2
Systematic Comparison of the Structures of Proteins from Mesophiles, Thermophiles, and Extreme Thermophiles

Property		Correlation with temperature	Change in proteins from moderately thermophiles	Change in proteins from extremely thermophiles
Cavities	number	↓↓	0	↓↓↓
	volume	↓	↑	↓
	area	↓	↑	↓↓
Hydrogen bonds	number	0	0	0
	un-satisfied	↓	↓	↓
Ion pairs	<4.0 Å	↑↑	↑	↑↑↑
	<6.0 Å	↑↑	↑↑	↑↑↑
	<8.0 Å	↑↑↑	↑↑↑	↑↑↑
Secondary structure	α	0	↑	0
	β	↑	0	↑↑
	irregular	↓	↓	↓
Polarity of surfaces	exposed	↓↓	↑↑↑	0
	buried	0	↑	↑

Note: The number of arrows (1, 2 or 3) shows whether the represented correlation or change is considered insignificant, moderately significant or highly significant.

fraction of residues (32.0%) in α-helices compared with mesophilic proteins (25.4%), and their helices seem to be more stable: the fraction of the helix-disfavoring residues His, Cys, and Pro residues in helices is lower in thermophilic proteins than in mesophilic proteins, whereas the fraction of the helix-favoring Arg residues is higher. On the level of the amino acid content, the proportions of the thermolabile residues Cys and Ser are significantly lower, while Arg and Tyr are more frequent in thermophilic proteins. In accordance with Szilagyi and Závodszky (2000), salt bridges and side chain-side chain hydrogen bonds are more frequent in thermophilic than in mesophilic proteins (Figure 3).

In a computational study, the number and the electrostatic energies of the salt bridges found in the subunits of the homohexameric glutamate dehydrogenases from the hyperthermophile *P. furiosus* (pfGDH; T_m = 113°C) and from the mesophile *Clostridium symbosium* (csGDH; T_m = 55°C) were compared (Kumar et al., 2000b). pfGDH and csGDH show an amino

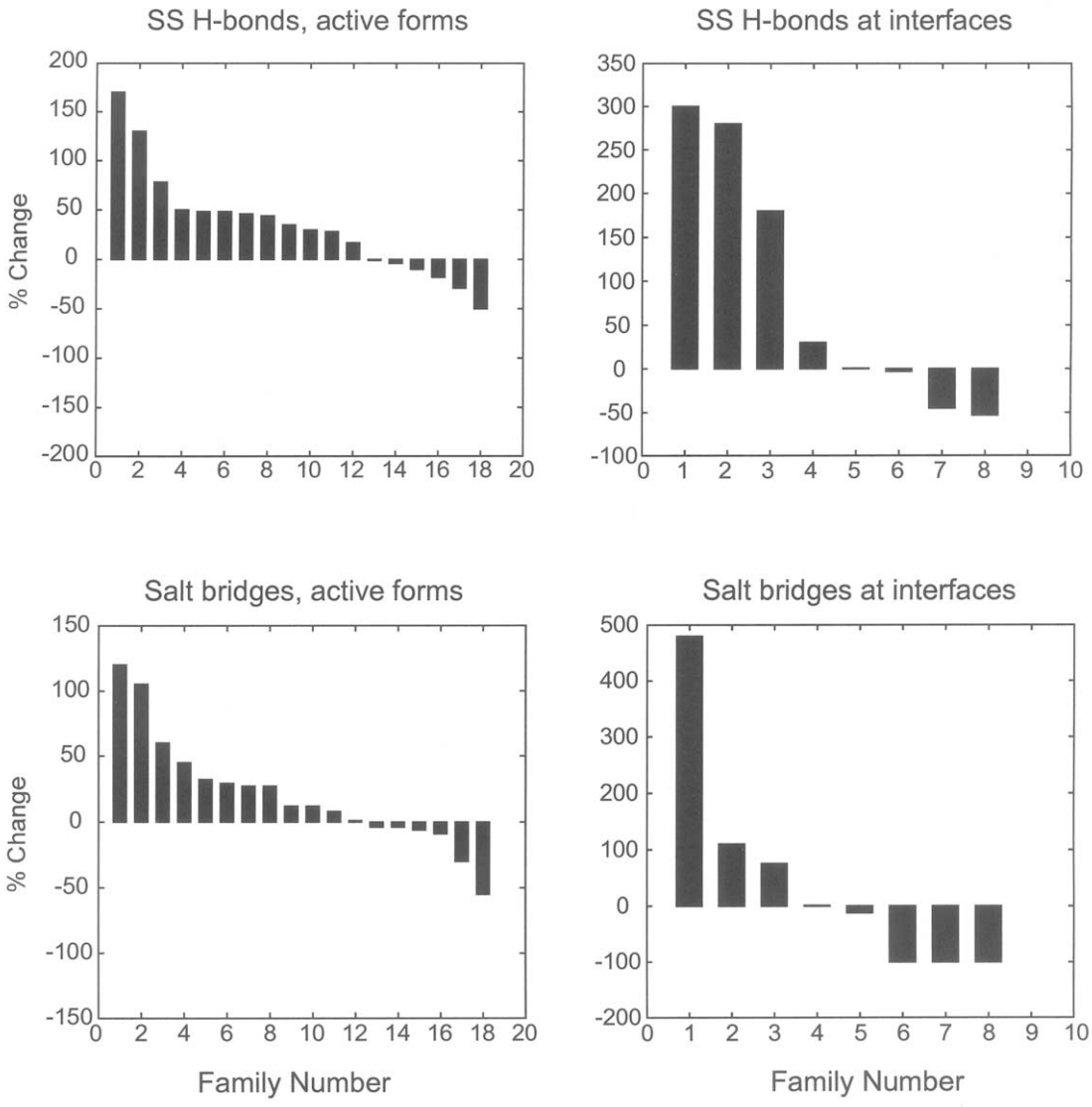


FIGURE 3. Thermophilic members of most investigated protein families contain an increased number of side chain-side chain hydrogen bonds (SS H-bonds) and salt bridges. The structurally most similar mesophilic-thermophilic protein pairs within each of the 18 investigated families were compared with respect to several properties. For each subplot, the numbers at the x-axis denote a particular protein family, while the y-axis represents the percent change in the property indicated at the top. A positive change indicates that the thermophilic family member has a higher content than the mesophilic family member. The N-O distance limit for salt bridge formation was 4 Å. The data on interfaces are available only for eight families. (Reprinted with permission from Kumar, S., Tsai, C.-J., and Nussinov, R. *Prot. Eng.* **13:** 179–191. Copyright (2000) Oxford University Press.)

acid sequence identity of 34%, and their backbones superimpose with an rms deviation of 1.38 Å. The N-O distance cut-off for the formation of a salt bridge was set to 4 Å. A salt bridge was classified as being "exposed" if the solvent accessible surface areas (ASA) of the involved residues X were (on average) above 20% of the ASA calculated for the tripeptide Gly-X-Gly in an extended conformation; otherwise, a salt bridge was classified as being "internal". The total electrostatic energy of a salt bridge, $\Delta\Delta G_{tot}$, was calculated relative to a computer mutation of the salt bridge side chains to their hydrophobic isosteres, using the following equation:

$$\Delta\Delta G_{tot} = \Delta\Delta G_{dslv} + \Delta\Delta G_{brd} + \Delta\Delta G_{prt} \quad (16)$$

where DDG_{dslv} represents the unfavorable desolvation free-energy penalty (desolvation penalty) incurred due to the desolvation of the salt bridge-forming side chains from water (unfolded state U) to the protein interior (folded state F), $\Delta\Delta G_{brd}$ represents the electrostatic free energy between two salt bridge-forming side chains in the folded protein, and $\Delta\Delta G_{prt}$ represents the interaction free energy of salt bridge-forming side chains with the rest of the protein. Each of the above terms were computed using continuum electrostatics, both for room temperature and for the respective optimum growth temperatures of *P. furiosus* (100°C) and *C. symbosium* (37°C). Detailed calculations were performed for 29 out of 40 intrasubunit salt bridges in pfGDH and for 17 out of the 20 intrasubunit salt bridges in csGDH monomers. At 100°C, only 2 out of 29 salt bridges of pfGDH are destabilizing. When taking into account the decrease of the dielectric constant of water with temperature (about 80 at 25°C, about 55 at 100°C), the average $\Delta\Delta G_{tot}$ amounts to –27 kJ/mol at 100°C. In csGDH, 6/17 salt bridges are destabilizing, and another six are only marginally stabilizing. As a result, the average $\Delta\Delta G_{tot}$ of a salt bridge in csGDH is only about –6.3 kJ/mol at 37°C. The reason for the difference in $\Delta\Delta G_{tot}$ between pfGDH and csGDH is that the desolvation penalty is hardly compensated by the $\Delta\Delta G_{brd}$ and $\Delta\Delta G_{prt}$ in csGDH, but easily accommodated in pfGDH. It appears that in pfGDH the interaction of the salt bridge-forming residues with the rest of the protein is almost as significant as is the electrostatic interaction. This might be due to the existence of many energetically favorable salt bridge networks in pfGDH, which comprise 23 out of 29 of the investigated salt bridges.

Karshikoff and Ladenstein (1998) concentrated on the role of the packing density in the stabilization of thermophilic proteins. To this end, the partial specific volumes, cavity volumes (space that is sufficiently large to accommodate at least one water molecule), and the smaller void volumes were calculated for a set of 80 nonhomologous monomeric mesophilic proteins with high-resolution X-ray structures and for a set of 24 proteins from thermophilic and hyperthermophilic organisms. No significant difference between mesophilic and thermophilic proteins was detected for all three parameters, and therefore it was concluded that packing density is not a dominant factor for high protein thermostability.

Two studies specifically investigated the intrinsic stability of α-helices in mesophilic and thermophilic proteins. Petukhov et al. (1997) compared the energy characteristics of α-helices from thermophilic and mesophilic proteins within four different protein families (RecA proteins, L-lactate dehydrogenases, seryl-tRNA synthetases, aspartate aminotransferases). The intrinsic free energy for the formation of homologous α-helical fragments (ΔG_{int}) was estimated with a modified model of the statistical me-

chanical theory of helix/coil transitions. This model considers all known factors that are important for the stability of α-helices, which is hydrogen bonding in the backbone (main chain-main chain hydrogen bonds), the intrinsic helical properties of amino acids (mainly to the loss of entropy required for adoption of helical dihedral angles), interactions between side chains ($i,i+3$; $i,i+4$), electrostatic interactions between charged side chains and the helix macrodipole, and the capping interactions at N- and C-termini, where charged and polar side chains can form additional hydrogen bonds with free NH and CO groups of the main chain. It was concluded from these calculations that thermophilic and mesophilic proteins differ in total helix stability rather than in the stability of individual helices, and that the main contributors to helix stability are the intrinsic helical propensities of amino acids and side chain-side chain interactions. Within the the RecA protein familiy, ΔG_{int} for the RecA proteins from *Pseudomonas aeruginosa*, *E.coli*, and *T. thermophilus* were equal at 18°C, 40°C, and 75°C, which are close to the corresponding growth temperatures of the organisms.

Facchiano et al. (1998) compared the X-ray structures of 13 thermophilic proteins with their mesophilic homologues. The comparison revealed no significant difference in helical content between mesophilic and thermophilic proteins. The stabilities of the individual helices were computed taking into account 10 different factors, which were similar to those used by Petukhov et al. (1997). In 69% of all cases, the helices of thermophilic proteins were estimated to be more stable, due to the higher intrinsic helical propensities of amino acids. This was mainly due to a lower content of the helix-destabilizing β-branched amino acids Val, Ile, and Thr.

Several studies focused on the structural comparison of iso-functional enzymes from psychrophilic, mesophilic, thermophilic, and hyperthermophilic sources. Maes et al. (1999) compared the X-ray structures of 10 different triosephosphate isomerases (TIM), 1 psychrophilic (from *Vibrio marinus*), 7 mesophilic, 1 moderate thermophilic (from *B. stearothermophilus*), and 1 hyperthermophilic (from *T. maritima*). With the exception of TIM from *T. maritima*, which is tetrameric, all other compared TIMs are dimeric. Stabilization seems to be due to an increase of the hydrophobic surface area that is buried after folding of the monomer (ΔASA), which — compared with the psychrophilic TIM — is increased by 3% in the moderately thermophilic and by 8% in the hyperthermophilic TIM. Furthermore, the hydrophobic area that is buried upon assembly of the monomers into oligomers is increased by 7% in the moderately thermophilic and, due to the higher association state, by 50% in the hyperther-mophilic TIM. With respect to salt bridges, the situation is less clear. Both the psychrophilic and the hyperthermophilic TIMs contain a higher number of salt bridges (0.11 and 0.09 per residue) compared with the average value of mesophilic TIMs (0.07 per residue). The increase in *T. maritima* TIM is mainly due to the four intersubunit salt bridge networks at the new dimer-dimer interface. In contrast, the moderately thermophilic TIM from *B. stearothermophilus* does not contain an increased number of salt bridges. In summary, it can be concluded that TIM from *T. maritima* is stabilized by improved hydrophobic interactions and a higher number of salt bridges, and that both features are due mainly to the increased association state.

Auerbach et al. (1998) determined the X-ray structure of lactate dehydrogenase (LDH) from *T. maritima* and compared it with five mesophilic and moderately thermophilic LDHs. All investigated LDHs are tetramers and no significant difference in the number of *inter*subunit salt bridges was

found. The extremely and moderately thermophilic LDHs contain, however, an increased number of *intra*subunit ion pairs: at an N-O distance cut-off of 4 Å, about 17 ion pairs per monomer of LDH from *T. maritima* and *B. stearothermophilus* are found, corresponding to 0.05 per amino acid residue. The three mesophilic and psychrophilic LDHs contain about 10 ion pairs per monomer corresponding to 0.03 per residue. Other stabilizing features seem to be a higher content of secondary structure and a decrease of the ratio of hydrophobic to charged surface area. Compared with LDHs from *B. stearothermophilus* and the mesophile *Bifidobacterium longum*, the subunits of LDH from *T. maritima* contains fewer cavities (two compared to four) and a significantly decreased total cavity volume (39.8 $Å^3$ compared with 109.6 $Å^3$ and 163.5 $Å^3$, respectively). Furthermore, the amount of intersubunit contacts is increased in the hyperthermophilic LDH, suggesting that the subunits are bound more tightly in thermophilic LDHs.

3. Summary

The results of the discussed comprehensive amino acid sequence and structural comparisons can be summarized as follows. Proteins from (hyper-) thermophiles are stabilized in various ways. Some stabilizing features seem to be of minor importance or to be crucial only in certain protein families. Among them are an increased number of hydrogens bonds, higher packing densities and secondary structure contents, improved hydrophobic interactions, optimized surface areas, decreased volumes, fewer cavities, and a shortening of the polypeptide chain. Those stabilizing features that have been detected in most studies in which they were systematically tested are (1) the stabilization of α-helices, (2) an increase in the fraction of Pro and β-branched amino acids, which decrease the entropy of the unfolded state, (3) a decrease in uncharged polar residues, which avoids chemical degradation of the polypeptide chain at the high physiological temperatures of thermophiles. The most general feature is (4) an increased number of electrostatic interactions and their optimization, for example, by the formation of salt-bridge networks, in thermophilic and hyperthermophilic proteins.

B. Mutational Analysis of Protein Thermostability

A large number of mutational studies have been performed in order to understand the role of individual residues in providing the stabilizing features that were discussed in the preceding sections. To this end, selected amino acids were exchanged by site-directed mutagenesis, and the resulting changes in stability were measured. Some recent and relevant case studies are discussed here.

1. Case Studies

Phosphoribosylanthranilate isomerase (PRAI) is a $(\beta\alpha)_8$-barrel enzyme of tryptophan biosynthesis, which is monomeric and labile in most mesophiles, but an extremely thermostable homodimer in *T. maritima*, with a half-life at 85°C of 310 min (Sterner et al., 1996). The two monomers are associated via the N-terminal faces of the central β-barrels (Figure 4), and the side chains of the N-terminal methionines and the C-terminal leucines of both subunits are buried in a hydrophobic cluster (Hennig et al., 1997). Most strikingly, two long, sym-

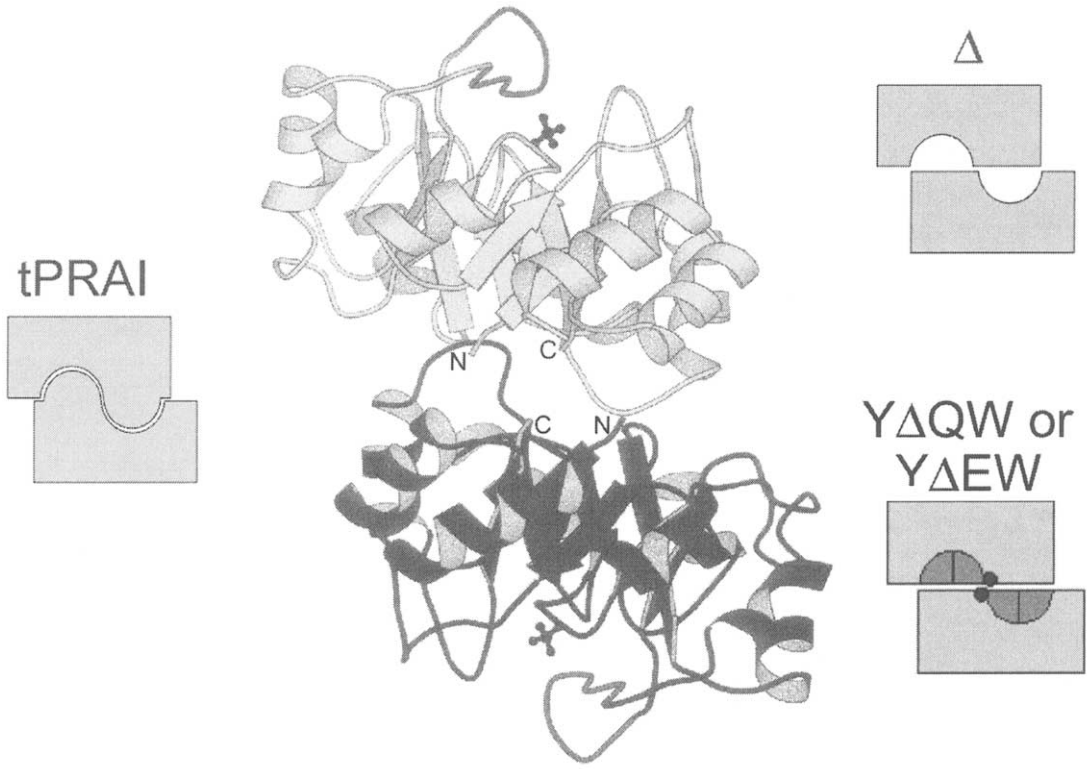


FIGURE 4. Ribbon diagram of the parental dimer of phosphoribosylanthranilate isomerase from *Thermotoga maritima* (tPRAI) viewed down the molecular twofold axis. The active sites are identified by the bound phosphate ions (tetrahedrons). The two identical $(\beta\alpha)_8$-barrel subunits associate through the N-terminal faces of their central β-barrels. The schematic insets represent intermediate steps in converting the dimer interface of tPRAI, with interdigidating surface loops and hydrophobic pockets, to that of the monomeric variants YΔQW or YΔEW. (Reprinted from *Structure* **8:** Thoma, R., Hennig, M., Sterner, R., and Kirschner, K. Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima*, pp. 265–276, Copyright (2000), with permission from Elsevier Science.)

metry-related loops that connect α-helix 2 with β-strand 3 protrude reciprocally into cavities of the β-barrel of the other subunit and provide for multiple hydrophobic interactions. In order to test the role of dimerization for the function and stability of PRAI from *T. maritima* (tPRAI), monomeric variants were generated by site-directed mutagenesis at the dimer interface (Thoma et al., 2000). A simulated "dimer" interface of monomeric PRAI from *E. coli* (ePRAI) was used as a guide for planning the amino acid exchanges. Deletion of residues Pro52 and Phe53, which are located at the tip of the loop α2β3, shortened the loop to the length found in ePRAI but left an energetically unfavorable cavity in the β-barrel of the adjacent subunit (change designated Δ in Figure 4). This cavity was filled by replacing two residues pointing into the interior of the β-barrel by the bulkier residues found at the equivalent positions in ePRAI: Ala25 was replaced by Tyr (change designated Y) and Ile101 by Trp (change designated W). Phe55 of one subunit and Phe55a of the second subunit are located close to each other at the twofold symmetry axis of the dimer. In order to weaken the reciprocal hydrophobic interactions, Phe 55 was replaced by Gln (change designated Q), which is the equivalent residue in ePRAI. Alternatively, Phe55 was replaced by Glu (change

designated E), introducing electrostatic repulsion between the two negative charges and thus promoting the dissociation of dimers further. All three investigated variants Δ, YΔQW, and YΔEW turned out to be catalytically as active as wild-type tPRAI, indicating that extensive modifications at the N-terminal face of the β-barrel are not transmitted to the active sites at the C-terminal face (Table 3). The analysis of the association states of the variants by analytical gel filtration and analytical ultracentrifugation showed that deletion of Pro52 and Phe 53 (variant Δ) is already sufficient to dissociate the dimer. Filling the resulting cavity by the mutations Ala25Tyr and Ile101Trp and introducing the mutation Phe55Gln (variant YΔQW) partially reversed this effect and resulted in an equilibrium between monomers and dimer. These results suggest that repacking permits novel interactions at the interface of YΔQW. In contrast, the replacement Phe55Glu (variant YΔEW) leads to a purely monomeric protein. As unfolding of tPRAI and its variants is irreversible, only kinetic stabilities could be determined. The corresponding half-lives at 85°C are 310 min for the wild-type enzyme and 3 to 5 min for all the variants Δ, YΔQW, and YΔEW (Table 4). The half-lives of dimeric tPRAI and monomeric YΔEW were independent of protein concentration. In contrast, the half-life of YΔQW, which exists in a monomer-dimer equilibrium, increased with increasing protein concentration. Therefore, dimerization, significantly stabilizes YΔQW, and, by inference, also the wild-type tPRAI, but does not affect the catalytic activity.

The archaeal histones HFoB (T_m = 75°C) from the mesophile *M. formicum* and HmfB (T_m = 114°C) from the hyperthermophile *M. fervidus* are dimeric enzymes that unfold reversibly according to Equation 15 and their thermodynamic stabilities ΔG_{FU} could be determined (Li et al., 1998). The structural basis for the different stabilities of HFoB and HMfB were elucidated by site-directed mutagenesis at all sites at which HFoB and HMfB differ (Li et al., 2000). The core of the histone dimers consists of two long, antiparallel α-helices, each belonging to one monomer (α-helix 2 and 2a). Nine out of the 15 differences between HfoB and HMfB are located in this region. At the center of the α2 – 2a interaction, Ile 31 and Met 35 are found in hyperthermophilic HMfB compared with Ala 31 and Lys 35 in mesophilic HFoB. When the residues at position 35 were mutually exchanged, the T_m value of HFoB-K35M was increased by 15°C and the T_m value HMfB-M35K was lowered by 19°C. Compared to HFoB, HMfB contains additional *intra*subunit ion pairs, namely, E33-R37 and E33a-R37a, and *inter*subunit ion pairs, namely D14-R27a and D14a-R27. The T_m value of the HFoB-E37R variant was increased by 4°C, apparently by replacing two repulsive E33(a)-E37(a) interactions with two attractive E33(a)-R37(a) interactions. Accordingly, the T_m value of the HMfB-R37E variant was lowered by 6°C. It was concluded that the higher T_m value of HMfB compared with HFoB is due to improved intermolecular hydrophobic interactions near the centre of the histone dimer core and due to additional ion pairs at the dimer interface.

The enzymes 3-isopropylmalate dehydrogenase from *T. thermophilus* (ttIPMDH; T_m ~ 87°C) and *E. coli* (eIPMDH; T_m ~ 70°C) constitute another pair of homologous, dimeric proteins with different thermostabilities. A comparison of the X-ray structures suggested that an increased number of intersubunit ion pairs in ttIPMDH is important for its high thermostability. For example, Glu190 in ttIPMDH forms an ion pair with Arg144a in the other subunit, which is further engaged in an ion pair with Glu142a, that is, an ion triad is formed. In eIPMDH Glu190 is replaced by Gln200,

TABLE 3
Steady-State Enzyme Kinetic Constants at 25°C of Phosphoribosylanthranilate Isomerase from *T. maritima* (tPRAI) and Its Variants

Proteins[1]	designation	k_{cat} (s^{-1})	K_M^{PRA} (μM)	k_{cat}/K_M^{PRA} ($\mu M^{-1}s^{-1}$)
tPRAI	tPRAI	3.7	0.28	13.3
tPRAI variants				
deletion 52-PF-53	Δ	3.5	0.30	11.7
A25Y/Δ/F55Q/I110W	YΔQW	3.1	0.26	11.9
A25Y/Δ/F55E/ I110W	YΔEW	3.8	0.34	11.2
ePRAI[2]	ePRAI	34.5	12.2	2.8

1 Sequence changes as depicted schematically in Figure 4.
2 Monofunctional PRAI variant from *E.coli*.

Reprinted from *Structure* 8: Thoma, R., Hennig, M., Sterner, R., and Kirschner, K. Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima*, pp 265-276, Copyright (2000), with permission from Elsevier Science.

TABLE 4
Dimer Dissociation Behavior and Thermostabilities of Phosphoribosylanthranilate Isomerase from *T. maritima* (tPRAI) and Its Variants.

Proteins[1]	M_{calc}[2] (kDa)	gel filtration (25°C) M_{app}[3] (kDa)	sedimentation equilibrium (20°C) M_{app}[4] (kDa)	sedimentation equilibrium (20°C) K_d[5] (μM)	thermal inactivation[6] (85°C) half-life (min)
tPRAI	23.04	30.2	49.6	$< 10^{-5}$	310
tPRAI variants					
Δ	22.80	17.0	24.2	$> 10^{2}$	3
H-YΔQW	25.10	21.7	-	5	5
H-YΔEW	25.10	17.0	24.0	$> 10^{3}$	5
ePRAI[7]	21.10	18.5	23.0	nd	nd

1 Sequence changes as depicted schematically in Figure 4.
2 Calculated molecular mass for the monomer.
3 Apparent molecular masses at 25°C and 1 μ*M* eluted total monomer concentration.
4 Apparent molecular masses of predominant species at 20°C and 4 μ*M* initial total monomer concentration.
5 Dissociation constant K_d.
6 Exponential decrease of V_{max} characterized by the half-time $t_{1/2}$ at 0.5 μ*M* total monomer concentration.
7 Monofunctional PRAI variant from *E.coli*.

Note: nd: not determined

Reprinted from *Structure* 8: Thoma, R., Hennig, M., Sterner, R., and Kirschner, K. Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima*, pp 265-276, Copyright (2000), with permission from Elsevier Science.

which cannot engage in ion pair formation. However, there are two positively charged residues, Lys150a and Arg152a, on the other subunit of eIPMDH that could form an ion pair with Gln200, if it carried a negative charge. Mirror image mutations were performed (Nemeth et al., 2000). In ttIPMDH Glu190 was replaced by Gln, and in eIPMDH Gln200 was replaced by Glu. Because unfolding of all produced variants was irreversible, only apparent T_m values could be measured by DSC and far UV circular dichroism. For all variants, the T_m value increased with increasing protein concentration. These results suggest that a subunit dissociation/association preequilibrium is rate limiting in the unfolding mechanism of IPMDHs, which probably follows Equation 15. The Glu190Gln variant of ttIPMDH is significantly destabilized, and the Gln200Glu variant of eIPMDH is significantly stabilized compared with the corresponding wild-type enzymes. The destabilization of ttIPMDH by the replacement of a charged by a polar residue, and the corresponding stabilization of eIPMDH by the mirror image mutation is probably due to the destruction and the formation of intersubunit ion pairs.

The effects of the removal of salt bridges was tested for two enzymes from the hyperthermophile *T. maritima*, glyceralaldehydephosphate dehydrogenases (GAPDH; Pappenberger et al., 1997), and indoleglycerol phosphate synthase (IGPS; Merz et al., 1999).

Both thermophilic and mesophilic GAPDHs are tetramers, but the investigated thermophilic variants contain a significantly increased number of intrasubunit ion pairs at the protein surface (Korndörfer et al., 1995; Tanner et al., 1996). GAPDH from *T. maritima* (tGAPDH) contains four charged side-chains (Arg20, Arg320, Asp 323, Glu326) that form an ionic network, which is reduced to the Arg320-Asp323 ion pair in GAPDHs from *B. stearothermophilus* and lobster. Arg20 is part of helix αB, which is located close to the N-terminus, and Arg320, Asp 323, Glu326 are part of helix α3, which is located close to the C-terminus of each subunit. The importance of the salt bridge cluster for the kinetic stability of tGAPDH was tested by generating the variants Arg20Ala and Arg20Asn (Pappenberger et al., 1997). Both variants were significantly destabilized compared with the wild-type enzyme: the half-lives at 100°C were decreased about four-fold, which translates into a decrease of the free energy of activation for thermal unfolding ($\Delta\Delta G_u^{\ddagger}$) of 4 kJ/mol at 100°C. Interestingly, the effects of the substitutions on the kinetic stability of tGAPDH are strongly temperature dependent. The stability differences are expected to disappear at temperatures beyond 105°C, but are very significant at 80°C ($\Delta\Delta G_u^{\ddagger}$= 10 kJ/mol), where *T. maritima* grows optimally.

IGPS is a monomeric $(\beta\alpha)_8$-barrel enzyme of tryptophan biosynthesis. The hyperthermophilic IGPSs from *T. maritima* (tIGPS, half-life = 11 min at 90°C) and *Sulfolobus solfataricus* (sIGPS, half-life = 15 min at 90°C) contain twice the number of potentially stabilizing salt bridges compared with the mesophilic IGPS from *E. coli* (eIGPS, half-life = 0.5 min at 55°C) (Hennig et al., 1995; Merz et al., 1999). The N-termini of both sIGPS and tIGPS are fixed to the surface of the protein by a salt bridge, which is missing in eIGPS. In tIGPS, this salt bridge is formed between Asp184 and Arg2, which is part of the extra N-terminal helix α_0. Several salt bridges in sIGPS connect adjacent α-helices of the $(\beta\alpha)_8$-barrel. In tIGPS, such a salt bridge (Glu73 - Arg241) crosslinks helices α1 and α8, providing a noncovalent clamp (Figure 5) across the closure of the barrel. In order to investigate the contribution of these salt bridges to the thermostability of tIGPS, the variants Arg241Ala and Asp184Ala were generated and their kinetic stabilities at 85.5°C were

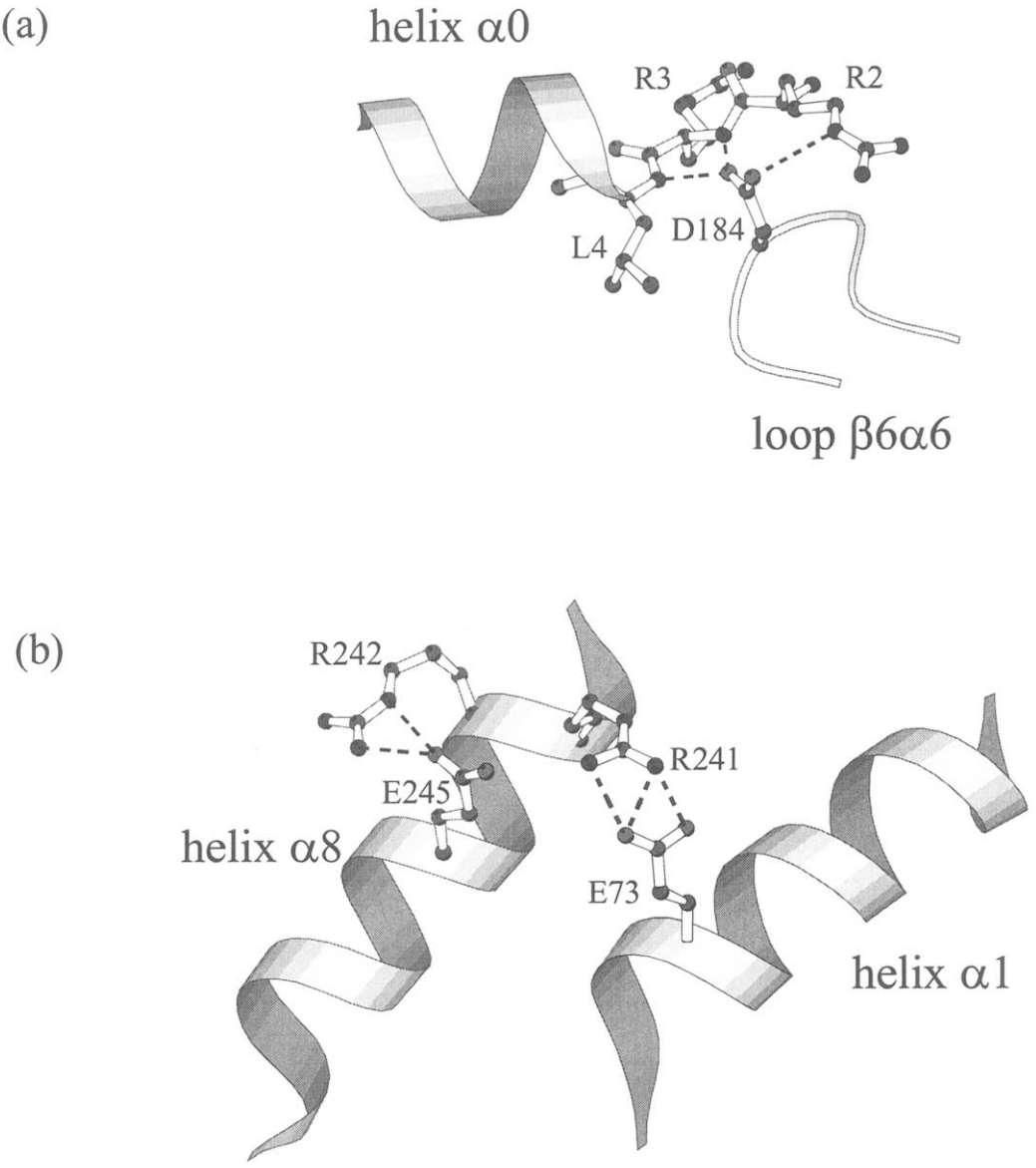


FIGURE 5. Potentially stabilizing salt bridges in indoleglycerolphosphate synthase from *Thermotoga maritima* (tIGPS). (a) Salt bridge (R2-D184) fixes the N-terminus of tIGPS to the $(\beta\alpha)_8$ barrel. Ribbon presentation of the backbone conformations of helix α_0 and loop $\beta_6\alpha_6$. The side-chains of R2, R3, L4 and D184 are depicted as ball-and-stick models. (b) Salt bridge (E73-R241) between helix α_1 and helix α_8. Ribbon presentation of the backbone conformations of the helices. The side-chains of E73, R241, R242, and E245 are depicted as ball-and-stick models. Residues R242 and E245 form an intrahelical salt bridge within helix α_8. The broken lines indicate salt bridges with interatomic distances $\leq$ 3.7 Å. (Reprinted with permission from Merz, A., Knöchel, T., Jansonius, J.N., and Kirschner, K. *J. Mol. Biol.* **288:** 753–763. Copyright (1999) Academic Press.)

compared with that of the wild-type enzyme (Merz et al., 1999). The half-lives were 42 min for the wild type, 36 min for the Asp184Ala, and 14 min for the Arg241Ala variants, showing that the two probed salt bridges stabilize tIGPS, albeit to different extents.

Strop and Mayo (2000) investigated by double mutant cycles the role of two salt bridges for the stability of an iron-free variant of rubredoxin from *Pyrococcus furiosus* (pRdx, cf. Section II.C and VI; Hiller et al., 1997; Cavagnero et al., 1995, 1998). The side chain to side chain salt bridge between Lys6 and Glu49 is not stabilizing. Because the two residues involved in a salt bridge are immobilized, which is entropically unfavorable, the favorable electrostatic interaction between the side chains of Lys6 and Glu49 might be compensated (Kumar et al., 2000b). In contrast, the main chain to side chain salt bridge between the positively charged N-terminus and Glu14 provides a significant thermodynamic stabilization energy. This salt bridges might be stabilizing, because the main chain is fixed in the folded protein and the entropic cost arises only for one (the side chain) partner of the salt bridge. In accordance with this hypothesis, the N-terminus of *P. furiosus* rubredoxin is involved in β-sheet formation.

2. Summary

The case of PRAI from *T. maritima* demonstrates that the noncovalent assembly of subunits can drastically stabilize hyperthermophilic proteins. In support of the stabilizing effect of oligomerization, the native formyltransferase from the hyperthermophile *Methanopyrus kandleri*, which is a tetramer, is significantly more thermostable than an engineered dimeric variant of the enzyme (Shima et al., 2000). Archaeal histones and IPMDH from *T. thermophilus* appear to be stabilized by the optimization of intersubunit interactions, which are predominantly hydrophobic or electrostatic in nature. Along these lines, the extremely high thermostability of homodimeric citrate synthase from *Pyrococcus furiosus* appears to be due to an intersubunit ionic network, which is absent in the less thermostable homologue from *Thermoplasma acido-philum* (Arnott et al., 2000). The studies on the hyperthermophilic GAPDH, IGPS, and Rdx confirm the crucial role of salt bridges for stabilization (Perutz and Raidt, 1975). These studies and investigations of glutamate dehydrogenases from various sources (Vetriani et al., 1998; Kumar et al., 2000b) also show that the stabilizing effect of a given salt bridge depends on its structural context. Salt bridges that connect N- and C-termini appear to contribute significantly to the stability of both tIGPS and tGAPDH, probably by preventing the "fraying" of the N- and C-termini, which might initiate thermal denaturation. In pRdx, only one of two probed salt bridges is stabilizing, probably because its formation does not further decrease the entropy of the polypeptide chain. Similar entropic arguments were put forward to support the idea that clusters of salt bridges (ion pair networks) are more stabilizing than individual salt bridges (Yip et al., 1995; Hennig et al., 1995).

The evidence that electrostatic interactions can play a central role in the stabilization of hyperthermophilic proteins disagrees with the results of previous theoretical studies and mutational analyses of model mesophilic proteins. The effects of introducing or deleting salt bridges in T4 lysozyme, barnase, and *Staphylococcus nuclease* varied considerably, but were often insignificant for stability or even destabilizing (Pace, 2000; Kumar et al., 2000b).

IV. TEMPERATURE DEPENDENCE AND SPECIFICITY OF ELECTROSTATIC INTERACTIONS IN PROTEINS

These findings raise the question, whether electrostatic interactions are more stabilizing at high than at low temperatures. The answer depends on the temperature dependence of the physical properties of water and of the protein-water interactions (Pace, 2000). According to Coulomb's law, the electrostatic interaction energy E that is necessary to bring two unit charges q_1 and q_2 from a distance r to an infinite distance is inversely proportional to the dielectric constant D of the medium: $E = q_1 \cdot q_2 / D \cdot r$. Because the dielectric constant of water drops from about 80 at 25°C to about 55 at 100°C, favorable electrostatic interactions within proteins will be stronger at higher temperatures. When a salt bridge is formed, the energetically favorable hydration shell is removed from the involved charged amino acid side chains. Theoretical studies by Elcock (1998), which take into account the temperature dependence of hydration free energies of the 20 naturally occurring amino acids (Elcock and McCammon, 1997), suggest that this desolvation penalty would decrease as temperature increases. In a related study, molecular dynamics simulations of the Sac7d protein from *Sulfolobus solfataricus* were performed at three different temperatures (de Bakker et al., 1999). It was concluded that electrostatic protein-solvent interactions become less favorable as temperature increases, due to the increased thermal motion of water. In contrast, electrostatic intraprotein interactions are assumed to become more favorable with increasing temperature, due to a tightening of salt bridges. In another theoretical study, four different hyperthermophilic proteins were compared with their mesophilic homologues, and it was shown in every case that electrostatic interactions are more favorable in the hyperthermophilc proteins (Xiao and Honig, 1999). Most strikingly, the authors suggest that the key stabilizing feature is neither the number of ion pairs nor the formation of ion pair networks, but the optimum placement of the charged groups on the protein surface. This hypothesis is supported by several experimental studies (Grimsley et al., 1999; Loladze et al., 1999; Perl et al., 2000; Spector et al., 2000), two of which are discussed here in some detail.

Loladze et al. (1999) investigated the thermodynamic stability of ubiquitin, a much-studied small and monomeric protein that unfolds according to the two-state model. Based on a simple model (Ibarra-Molero et al., 1999), an energy value W_i for the interaction of each charged residue with the rest of the ionizable groups on the surface of ubiquitin was calculated. A positive value of W_i for a given charged group indicates that it destabilizes the protein electrostatically, suggesting that its neutralization or reversal increases the stability of ubiquitin. When W_i adopts a negative value, its replacement should destabilize ubiquitin (Figure 6). On the basis of these predictions, a number of amino acids were exchanged individually, and the stabilities of the corresponding ubiquitin variants were measured. Compared with the wild type, both stabilized and destabilized variants were generated, and there was good qualitative agreement between prediction and experiment. These experiments demonstrate that a protein can be stabilized by the optimization of long-range electrostatic interactions at its surface.

The cold shock proteins from the mesophile *B. subtilis* (*Bs*-CspB; T_m = 50°C) and the thermophile *B. caldolyticus* (*Bc*-Csp; T_m = 72°C) are reversibly unfolding proteins (Perl et al., 1998) with very similar amino acid sequences

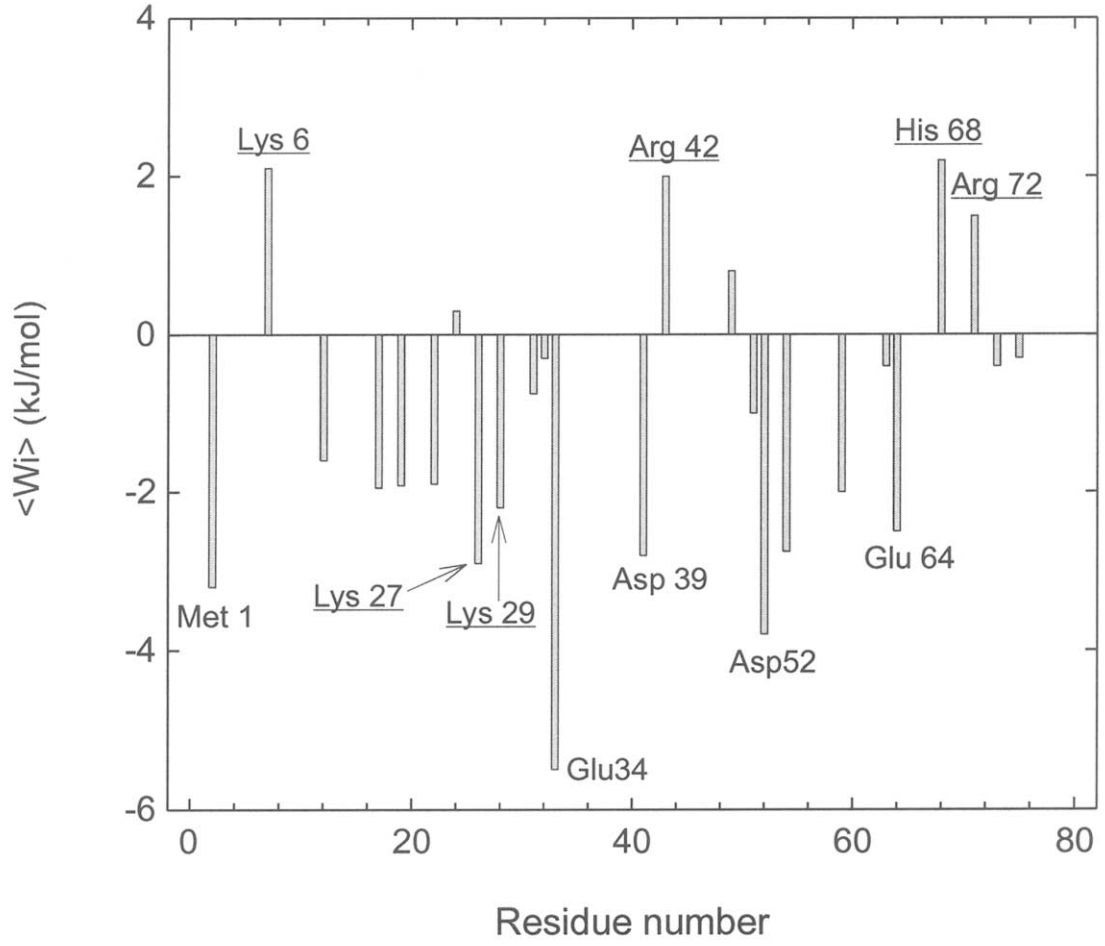


FIGURE 6. Energies due to the electrostatic interactions of all ionizable residues in the ubiquitin molecule at pH 5.0. Positive/negative values of $<W_i>$ predict that the amino acid side chains are involved in predominantly destabilizing/stabilizing charge-charge interactions. Underlined amino acids were exchanged experimentally. (Reprinted with permission from Loladze, V.V., Ibarro-Molero, B., Sanchez-Ruiz, J.M., and Makhatadze, G.I. *Biochemistry* **38:** 16419–16423. Copyright (1999) American Chemical Society.)

(55 out of 67 residues are identical) and virtually identical central β-sheets and binding regions for nucleic acids (Mueller et al., 2000; cf. Section II.C). The difference in ΔG_{FU} (at 70°C) between *Bc*-Csp and *Bs*-CspB amounts to about 16 kJ/mol in the absence of salt and to about 8 kJ/mol in the presence of 2 *M* NaCl, showing that differential electrostatic interactions contribute significantly to the observed stability difference (Mueller et al., 2000). All 12 amino acid exchanges between *Bc*-Csp and *Bs*-CspB are located at the protein surface. In order to elucidate the effect of each of these differences, the corresponding amino acids were mutually exchanged between the two proteins (Perl et al., 2000). From the 12 *Bc*-Bsp single mutants that contained the corresponding amino acids of *Bs*-CspB, only the Leu66Glu and Arg3Glu variants were destabilized. The double mutant Leu66Glu + Arg3Glu was slightly less stable than *Bs*-Csp. Accordingly, the corresponding *Bs*-CspB double mutant Glu3Arg + Glu66Leu was almost as stable as wild-type *Bc*-Csp (Figure 7). The analysis of the salt dependence of ΔG_{FU} showed that the Arg to Glu difference at position 3 fully accounts for the strong electrostatic contribution to the stability difference between *Bc*-Csp and *Bs*-CspB.

Is the stabilizing electrostatic effect of Arg3 local or global? Inspection of the crystal structure of *Bc*-Csp makes the formation of an ion pair between Arg3 and Glu46 unlikely. In accordance with this view, the *Bc*-Csp-Glu46Ala variant is not significantly destabilized compared with wild-type *Bc*-Csp (Figure 7). The formation of a stabilizing ion pair between Arg3 and the C-terminus of *Bc*-Csp can also be excluded, because the C-terminal extension of *Bc*-Csp by one amino acid (to reach the length of *Bs*-Csp) does not affect ΔG_{FU} (Figure 7). Because no other negatively charged side chain approaches Arg3 closer than 6 Å, the electrostatic stabilization caused by Arg3 does not seem to involve specific ion pairs, but to improve the global electrostatic potential of

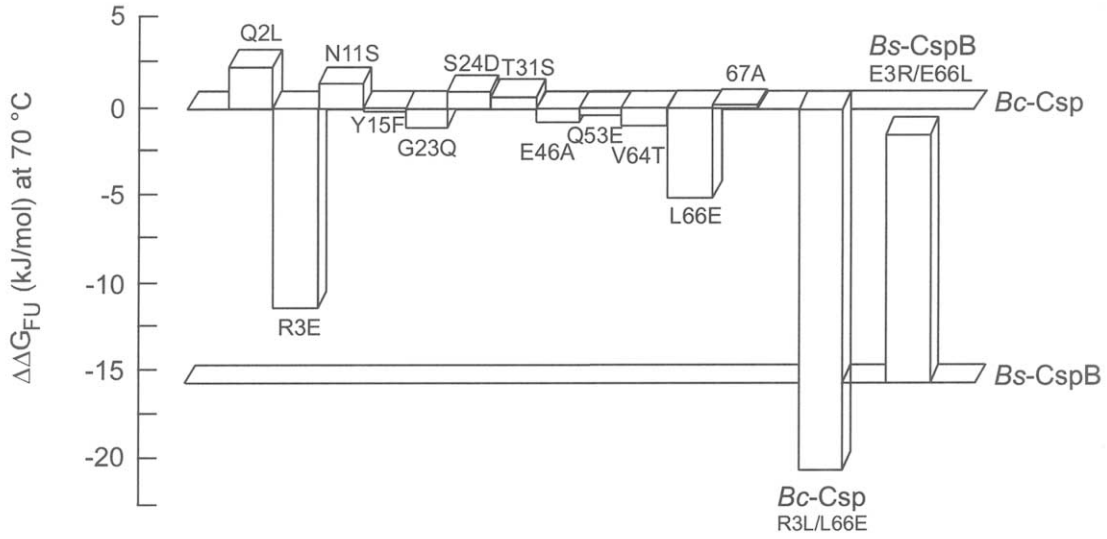


FIGURE 7. Effect of amino acid exchanges on the stability of *Bc*-Csp and *Bs*-CspB. The relative stabilities of the parent proteins are indicated by the horizontal lines. The first 12 bars display the relative differences in stability ($\Delta\Delta G_{FU}$) caused by the individual exchange of residues of *Bc*-Csp with that of *Bs*-CspB. The two bars at the right show the destabilization of *Bc*-Csp and the stabilization of *Bs*-CspB, respectively, by the reciprocal double amino acid exchanges at positions 3 and 66. (Reprinted with permission from Perl, D., Mueller, U., Heinemann, U., and Schmid, F.X., *Nature Struct. Biol.* **7:** 380-383 [2000].)

the protein. What is the basis of the stabilizing effect of Leu66, which is the C-terminal residue of *Bc*-Csp? The side chain of Leu66 packs on the main chain hydrogen bonds between Val47, Ser48, and Val64, which link the beginning of β-strand 4 with the end of β-strand 5. Leu66 thus decreases the polarity around these hydrogen bonds, which are strengthened due to the exclusion of water as potential hydrogen bond competitor.

In summary, this set of experiments convincingly demonstrates that two surface-exposed residues are responsible for the large difference in thermodynamic stabilities between *Bc*-Csp and *Bs*-CspB. Roughly two-thirds of this ΔG_{FU} difference originates from the predominantly electrostatic contribution of Arg3, and about one-third results from hydrophobic interactions of Leu66 at the carboxy terminus.

V. THERMOSTABILIZATION OF PROTEINS

A. Rational Design

The foregoing sections have shown how sequence and structural comparisons, in combination with site-directed mutagenesis, can identify putative molecular determinants of protein thermostability. One of the long-term goals of such studies is to design proteins with desired stabilities (and activities) (Hough and Danson, 1999). Although computer algorithms that reliably predict amino acid exchanges to stabilize any given protein are not yet available, the rational design of thermostable proteins has been successful in several cases.

An encouraging example for rational protein design is the stabilization of the monomeric β1 domain of the Streptococcal protein G (Gβ1), a ββaββ protein (Malakauskas and Mayo, 1998). A computer algorithm, which fixes the protein backbone and generates side chain rotamer libraries, models several stabilizing factors simultaneously. First, pair-wise interaction energies between amino acid side chains and between the side chains and the backbone are calculated according to empirical energy functions. The resulting energies serve as input for the optimization portion of the algorithm, which selects both those side chains and their corresponding conformations that yield the global minimum energy of the system. The study focused on five out of 13 so-called boundary residues, the side chains of which lie between the buried core and the solvent-exposed surface and are located on elements of secondary structure. The starting point was a Gβ1 variant, the core of which was already optimized by the introduction of the three amino acid exchanges. For the five boundary residues under consideration, both amino acid identity and side chain conformation were optimized, allowing 16 out of 20 amino acids at each position. (Pro, Cys, Gly, Met were omitted in order to avoid possible disruption of secondary structure, formation of S-S bonds, and to limit the loss of side chain entropy.) As a consequence, a diversity of 10^6 sequences was created from which the energetically most favorable one was selected. It contains the three core mutations mentioned above and the following four boundary mutations: Thr16Ile, Thr18Ile, Thr25Glu, and Val29Ile. The Trp residue at position 43 retained its wild-type identity during the optimization procedure. Several variants with different combinations of the amino acid exchanges were generated experimentally and tested for their stability. The variant with all seven amino acid exchanges (Gβ1-c3b4) shows, compared with wild type, an increase in T_m by more than 16°C and an increase in ΔG_{FU} (at 50°C) by about 18 kJ/mol (Table 5). How do the in-

TABLE 5
Experimental thermodynamic data of the β1 Domain of Streptococcal Protein G (Gβ1) and Its Stabilized Variants.

Protein	T_m[1] (°C)	ΔG_{FU}[2] (kJ mol^{-1})	$\Delta\Delta G_{FU}$[3] (kJ mol^{-1})	$\Delta\Delta A$[4] (Å^2)
Gß1	83	11.7	0.0	0.0
Gß1-c3	91	13.8	2.1	106.1
Gß1-c3b1	93	17.1	5.4	100.9
Gß1-3cb2	>99	25.5	13.8	281.9
Gß1-c3b4	>99	29.7	18.0	301.6

1 Midpoint of thermal unfolding transition.
2 Free energy of unfolding at 50ºC.
3 Change in free energy of unfolding with respect to Gß1.
4 Change in hydrophobic surface area burial with respect to Gß1 calculated using modeled structures.

Reprinted with permission from Malakauskas, S.M. and Mayo, S.L. *Nature Struct. Biol.* **5:** 470–475 (1998).

troduced amino acid exchanges improve the stability of Gβ1? The "boundary" exchange Thr25Glu appears to stabilize the protein in various ways. Position 25 is close to the N-terminus of the α-helix and the negative charge of Glu might interact with the positive end of the helix macrodipole. Also, Glu has a higher helix propensity than Thr. Furthermore, the structural model as determined by NMR predicts a hydrogen bond between Glu25 and the backbone amide of Val21, which is located in the turn between the second β-strand and the helix. The other boundary mutations, all of which are Ile, appear to stabilize through increased burial of hydrophobic surface area. A redesign of six helix surface positions of Gβ1, which was based on the same computer algorithm, increased the T_m value by 6°C, probably due to increased helix propensities and more favorable helix dipole interactions (Strop et al., 2000).

Van den Burg et al. (1998) stabilized a moderately thermostable, monomeric metallo-protease from *B. stearothermophilus* by a combination of rational design (introduction of a Pro residue and a disulfide bridge) and the replacement of several amino acids with those of the closely related but more thermostable protease thermolysin. The half-life of the resulting eightfold mutant at 100°C was increased about 340-fold compared with the wild-type enzyme, making it even more stable than thermolysin. Possible reasons for this increased stability are the reduction of the entropy of the unfolded state by the introduction of the disulfide bridge, and glycine to alanine (or serine), as well as alanine to proline exchanges. At 37°C, the catalytic activities of the mutant and the wild-type enzymes were identical. This study shows that few amino acid exchanges are sufficient to generate a hyperthermophilic protein, without loosing high catalytic activity at low temperatures. Enzymes with this combination of properties are expected to be very useful for industrial applications (Section VIII).

B. Directed Evolution

Despite the few successful examples presented above, it is still difficult to elucidate the structural basis of high protein ther-

mostability and to design proteins with increased stabilities. A promising and relatively novel approach to achieve this goal is directed molecular evolution. It mimics the natural evolution process by using iterative rounds of random mutagenesis and selection in order to generate proteins with desired properties. Given the existence of an appropriate screening or selection system, the new variants may even acquire combinations of properties that are not encountered in proteins from naturally existing organisms, but are optimal for industrial applications (Arnold and Volkov, 1999). Recent technical advances such as DNA shuffling (Stemmer, 1994a,b) or the staggered extension process (Zhao et al., 1998) allow the recombination of beneficial mutations that are created in the first rounds of directed evolution. In order to learn generally applicable rules for rational protein design, the optimized protein variants are purified, the generated changes are quantified, and their molecular basis is elucidated. The great advantage of the directed molecular evolution compared with the "rational" design approach is that new properties can be generated without a detailed *a priori* knowledge of the structure of the protein under study. The following selected examples illustrate that proteins can be made more stable or catalytically more active, or both, by directed evolution.

Giver et al. (1998) stabilized *p*-nitrobenzyl esterase from *Bacillus subtilis* (pNBE) by a combination of random mutagenesis, DNA shuffling, and screening. pNBE is a monomer that consists of 490 amino acid residues. The wild-type gene was amplified by error-prone polymerase chain reaction (PCR), and *E. coli* cells were transformed with the resulting plasmid gene library and plated. The expression of the pNBE variants was induced and their enzymatic activities (A) were determined at room temperature, both before (A_i) and after (A_r) heating. Variants exhibiting a combination of high activity (A_i) and high thermostability (A_r/A_i) were purified and their thermostabilities were measured. The gene encoding the variant with the highest T_m value was used to parent the next generation, and this procedure of random mutagenesis and screening (of 500 to 2000 clones per generation) was repeated several times. Genes encoding the best variants from the fifth-generation library were recombined using DNA shuffling. The most thermostable of the resulting variants (6sF9) contains nine amino acid exchanges and has a T_m value that is 14°C above that of the wild-type enzyme. At room temperature, this variant has a somewhat higher catalytic efficiency than the wild type enzyme, which shows that high thermostability and high activity at room temperature are not mutually exclusive properties (Figure 8). Spiller et al. (1999) determined the X-ray structure of a modified 6sF9 variant that contained additional four stabilizing exchanges. The resulting variant has a sequence identity to wild type of 97%, but a T_m value that is increased by 17°C. The structural analysis revealed that the individual amino acid exchanges resulted in altered core packing, helix stabilization, introduction of surface salt bridges, and reduction of flexibility in surface loops.

A similar approach was followed to stabilize the psychrophilic subtilisin S41 from the Antarctic *Bacillus* TA41, which is a monomeric and thermolabile serine protease that consists of about 300 amino acids (Miyazaki and Arnold, 1999; Miyazaki et al., 2000). Because the half-life of S41 is highly Ca^{2+}-dependent, S41 mutant gene libraries were screened at low Ca^{2+} concentrations, in order to stabilize S41 by increasing the Ca^{2+} affinity. Thermostable S41 variants were identified in a similar way as described above. Beneficial mutations identified in the first screening round were com-

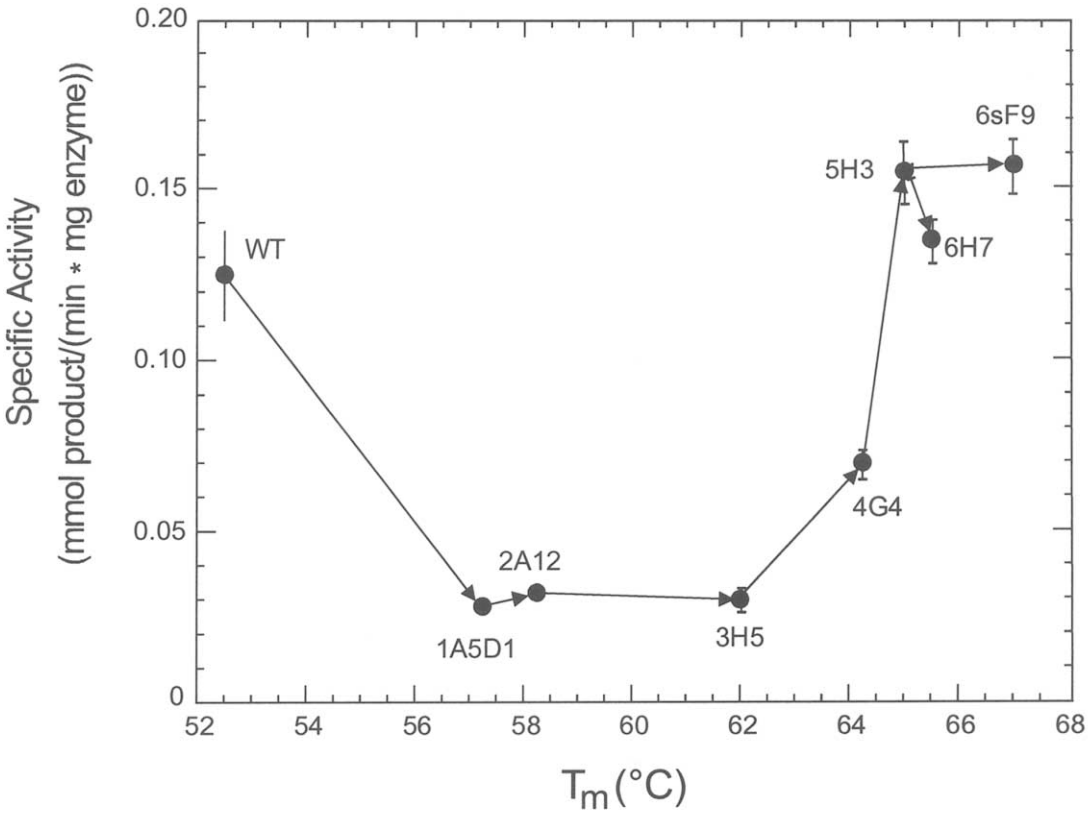


FIGURE 8. Specific enzymatic activities of purified wild-type (WT) and evolved *p*-nitrobenzyl esterases from *Bacillus subtilis* as a function of the melting temperature T_m. (Reprinted with permission from Giver, L., Gerhenson, A., Freskgard, P.-O., and Arnold, F.H. *Proc. Natl. Acad. Sci. USA* **95:** 12809-12813. Copyright (1998) National Academy of Sciences, U.S.A.)

bined and the resulting variants were further optimized. The most thermostable variant (3-2G7) contained seven amino acid exchanges. The stability differences between 3-2G7 and wild-type S41 were almost ten times greater at 1 m*M* Ca^{2+} than at 10 m*M* Ca^{2+}, due to an almost 100-fold increased Ca^{2+} affinity of 3-2G7. Enzyme activity measurements showed that 3-2G7 combines high thermostability with high activity at low temperatures: k_{cat}/K_M far surpassed that of wild-type S41 in the temperature range from 10 to 60°C, primarily due to an increase of k_{cat}. Ca^{2+} concentrations did not change the activity of either enzyme, confirming that Ca^{2+} binding increases stability independently of activity, that is, tighter Ca^{2+} binding does not interfere either with the active site geometry or with any protein motions required for catalysis. A comparison of the sequence of 3-2G7 with the sequences of mesophilic homologues S39 and SSII showed that none of the stabilizing amino acids is found at the corresponding positions in S39 and SSII. It can be concluded that the very large sequence space probably contains many different stabilizing amino acid substitutions, and that different ones were recruited by nature and by the directed evolution experiment.

The hitherto presented work uses *screening* of a large number of variants to identify proteins with increased stabilities. In contrast, Akanuma et al. (1998) applied *selection* in the extremely thermophilic host *Thermus thermophilus*, which can be naturally transformed with foreign DNA. IPMDH (encoded by leuB) from *B. subtilis*

(bsIPMDH) and *T. thermophilus* (ttIPMDH) are homodimers with a 56% identity in amino acid sequence but a 40°C difference in thermostability. The growth at elevated temperatures of *T. thermophilus* was coupled to the generation of stabilized bsIPMDH variants as follows. The bs-leuB gene was integrated into the chromosome of a leuB-deficient *T. thermophilus* strain. The resulting strain BTH 5601 grew at 56°C on plates lacking leucine, but not at 61°C. A large number of BTH5601 cells were plated on selective medium and incubated at 61°C. Transformants that were able to grow at this temperature presumably had acquired spontaneous mutations in bs-leuB that led to the stabilization of bsIPMDH. Two further rounds of selection were performed at 66°C and at 70°C, and a stepwise amino acid exchange with increasing selection temperature was found: Thr308Ile (selection at 61°C), Thr308Ile + Ile95Leu (selection at 66°C) and Thr308Ile + Ile95Leu + Met292Ile (selection at 70°C). The half-inactivation temperatures (cf. Section II.B) of the single, double, and triple mutants were increased by 4.5°C, 8.0°C, and 10.0°C, respectively. Inspection of the X-ray structure of ttIPMDH suggests that thermostabilization is caused by improved hydrophobic interactions and the removal of steric repulsion.

The stabilization that can be achieved by this approach is limited due to the rare occurrence of spontaneous mutations. In order to surmount this shortcoming, a combination of directed evolution and rational design was applied (Akanuma et al., 1999). Starting from the described triple mutant, bsIPMDH was further stabilized using a combination of random mutagenesis by error-prone PCR and selection in *T. thermophilus* at 73°C. This led to the acquisition of the additional exchanges Thr22Lys and Met256Val, and the half-inactivation temperature of the resulting quintuple variant was increased by another 3°C, compared with the triple mutant. Thr22Lys probably stabilized an α-helix, and Met256Val improved the hydrophobic interactions between the two subunits of bsIPMDH. On the basis of the X-ray structure, additional amino acid exchanges were introduced to further improve hydrophobic interactions between the two subunits of bsIPMDH (Glu253Leu) and to rigidify a turn structure (Glu112Pro, Ser113Gly, Ser115Glu). The resulting variant contained nine amino acid exchanges and its half-inactivation temperature was increased by 23°C compared with the wild-type bsIPMDH.

In a similar approach, the thermostability of the monomeric kanamycin nucleotidlytransferase (KNT) from *Staphylococcus aureus* was improved by random mutagenesis, DNA shuffling, and selection in *T. thermophilus* (Hoseki et al., 1999). *T. thermophilus* cells were transformed with a plasmid library containing randomly mutagenized KNT genes, plated on medium containing kanamycin and incubated at 64°C. A mixture of plasmids was prepared from the grown colonies, and the mutant genes were recombined using DNA shuffling. Further rounds of selection were performed at 69°C, 79°C, and 81°C. Each of the selected 10 most thermostable variants (KT3s) contained about 10 amino acids exchanges. One substitution, Val75Ala, was found in all ten cases and five out of 29 found exchanges are those to Pro. In order to generate a further stabilized variant, exchanges that were found in more than two other KT3s and all exchanges to Pro were individually incorporated into KT3-11. Nine of the introduced exchanges increased the thermostability of KT3-11. These exchanges were added simultaneously into KT3-11 to yield HTK, which contains 19 amino acid exchanges. The measurement of the irreversible thermal denaturation of wild-type KNT, KT3-11, and HTK yielded T_m values of 61°C, 73°C, and 84°C. The gene encoding HTK was

integrated into a plasmid, which can now be used for selection experiments at the upper growth limit of *T. thermophilus*.

DNA *family* shuffling recombines naturally occurring homologous genes (Crameri et al., 1998). This technique was used to generate chimaeras of the catalytic domains from a mesophilic (from *Streptomyces lividans*, SlxB-cat) and a thermophilic (from *Thermomonospora fusca*, TfxA-cat) pair of xylanases (Shibuya et al., 2000). SlxB-cat and TfxA-cat share an amino acid sequence identity of 72%, but their thermostabilities are drastically different. DNA from both genes was digested and recombined to yield a library of chimaeric xylanase genes. *E. coli* cells were transformed with this library, plated on RBB-xylan plates and incubated at 60°C for 1 h, which are conditions at which SlxB-cat is thermally denatured. Fifteen different thermally stable chimaeric enzymes were identified by halo formation. All of them possessed N-terminal regions derived from TfxA-cat, indicating that this segment was essential for activity under the selected conditions. These experiments show how directed evolution can be used to identify those segments of a proteins, if any exist, that are crucial for its thermostability.

Directed evolution has been used to improve the weak catalytic activities of hyperthermophilic enzymes at low temperatures. Indoleglycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus* (sTrpC) is $(\beta\alpha)_8$-barrel enzyme with a very low activity at 25° and 37°C compared with the corresponding enzyme from *E. coli* (eTrpC) (Hennig et al., 1995). The weak catalytic activity of sTrpC was increased by a combination of random mutagenesis and selection *in vivo* (Merz et al., 2000). An *E. coli* strain lacking the gene for *trp*C (Δ*trp*C) was transformed with a pool of randomly mutagenised plasmid-encoded s*trp*C genes, plated on minimal medium containing low concentrations of tryptophan and incubated at 37°C. Colonies growing faster than controls with wild-type *strp*C presumably contained sTrpC variants with an increased activity at 37°C. Inspection of the X-ray structure of sTrpC (Hennig et al., 1995) allowed the location of amino acid substitutions in faster growers. Most exchanges are located in α-helices and in loops that connect β-strands with α-helices at the C-terminal (active site) face of the barrel, but none of the previously identified catalytically essential residues of TrpC (Darimont et al., 1998) was altered. In about 40% of the sequence (between β-strands 2 and 5) no amino acid substitutions were found (Figure 9). The activity, flexibility and thermostability of several purified sTrpC variants was investigated. Compared with wild-type sTrpC, isolated variants have 2- to 4-fold increased k_{cat} values at 37°C, but drastically decreased (8- to 200-fold) K_M values (Table 6). As a consequence, the catalytic efficiencies (k_{cat}/K_M) are decreased 3- to 60-fold. These results confirm that k_{cat} is the selected trait and suggest that the intracellular CdRP concentrations are saturating for all variants.

What is the mechanistic basis of the increased k_{cat} values? In order to answer this question, it was first investigated which individual rate constant determines k_{cat} of sTrpC at 37°C. The minimal catalytic mechanism of the TrpC reaction (Figure 10) offers two possibilities: (1) the chemical step (ES to EP) is rate limiting: $k_3 << k_5$, and thus $k_{cat} = k_3$; (2) release of the product IGP (EP to E+P) is rate limiting: $k_3 >> k_5$, and thus $k_{cat} = k_5$. In order to distinguish between these possibilities, the rate constants of IGP release from the enzyme (k_5) and of IGP binding to the enzyme (k_6) were determined at 37°C for wild-type sTrpC and the F246S variant, using stopped-flow measurements under pseudo-first-order conditions ([CdRP]>>[sTrpC]). These measurements showed that k_5 (0.19 s^{-1}) is identical

FIGURE 9. Location of activating amino acid exchanges in the structure of indoleglycerolphosphate synthase from *Sulfolobus solfataricus* (sIGPS). H: α-helices, S: β-strands. Observed amino acid exchanges in activated variants are indicated below the sequence of sIGPS. The catalytically essential amino acids are underlined. The single cleavage site of trypsin is located C-terminal of Arg28, which is shown as an enlarged and bold type "R".

TABLE 6
Steady-State Enzyme Kinetic Constants of Indoleglycerol Phosphate Synthase from *S. solfataricus* (sIGPS) and Its Activated Variants.

protein	T (°C)	k_{cat} (s^{-1})	K_M (μM)	K_P (μM)	k_{cat}/K_M ($\mu M^{-1}s^{-1}$)
sIGPS	25	0.03	0.04	0.02	0.7
	37	0.15	0.05	0.03	3.0
	60	0.98	0.05	0.06	14.2
P2S	37	0.16	0.05	0.04	3.20
F246S	37	0.35	0.40	0.30	0.90
G212E	37	0.36	5.60	2.60	0.10
P2S/F246S	37	0.27	0.42	0.39	0.60
P2S/G212E	37	0.57	10.9	3.60	0.1
eIGPS[1]	25	3.60	0.42	0.54	8.60

[1] monofunctional IGPS variant of *E.coli*.

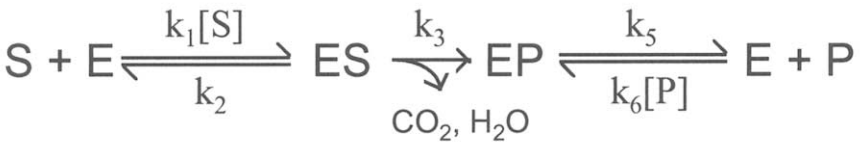


FIGURE 10. Minimal mechanism of the reaction catalyzed by sIGPS (E). S is 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate (CdRP). P is indolglycerol phosphate (IGP). $k_4 = 0$, because decarboxylation of CdRP renders the conversion of ES $\rightarrow$ EP irreversible. (Reprinted with permission from Merz, A., Yee, M. C., Szadkowski, H., Pappenberger, G., Crameri, A., Stemmer, W. P., Yanofsky, C., and Kirschner, K. *Biochemistry* **39:** 880–889. Copyright (2000) American Chemical Society.)

within the error limits to k_{cat} (0.15 s^{-1}), identifying product release as the rate-determining step of the sTrpC reaction at 37°C. Analogous measurements with the F246S variant yielded a much larger value for k_5 (9.6 s^{-1}), which equals about 25 times k_{cat} (0.35 s^{-1}). Obviously, product dissociation has become much faster, and k_{cat} probably equals the chemical step (k_3), which has become rate determining. In wild-type sTrpC, k_3 is probably also only 0.35 s^{-1} and thus only slightly larger than k_5. The effect of increasing temperature on k_{cat}, K_M and K_p suggests that the constipation by product release, which limits the activity of the parental enzyme at 37°C, is relieved at 60°C by a similar increase in flexibility as reflected by single- or double-residue replacements at 37°C.

In order to test the effect of the increased enzymatic activity on protein flexibility and stability, selected variants were digested with protease and irreversibly inactivated by heat. Trypsin cleaves solely at the R26-Q27 peptide bond of sTrpC (Darimont, 1994), which is located in a dynamically fluctuating loop that connects the additional α-helices α_0 and α_{00} (Figure 9). The half-lives of the full-length protein chains of several activated variants were decreased compared with wild type sTrpC, as estimated from SDS-PAGE. This is probably due to long-range perturbations that increase the flexibility of the loop between α_0 and α_{00}. Accordingly, binding of IGP significantly stabilized the most susceptible double-site variants P2S/F246S and P2S/G212E. At 87°C, the kinetic stability of the variants F246S and P2S/F246S is far lower than that of wild-type sTrpC, whereas the other variants are only slightly more thermolabile. A plausible reason for the labilizing effect of F246S is its location close to the C-terminus of sTrpC, at the end of helix α_8 (Figure 9). The phenyl ring of F246 is part of a hydrophobic cluster, which strengthens the interactions between modules $(\beta\alpha)_1$ and $(\beta\alpha)_8$, the noncovalent closure of the $(\beta\alpha)_8$ barrel. Substitution of the bulky, hydrophobic side chain by the small, hydrophilic side chain of Ser will weaken this cluster, resulting in a less stable protein.

In a similar experiment, the low-temperature activity of a tetrameric β-glucosidase from *Pyrococcus furiosus* was increased by directed evolution without affecting its extreme thermostability (Lebbink et al., 2000).

C. Summary

Due to the still limited understanding of the structural basis of protein (thermo-)stability, rational design of stabilized proteins has been successful only in few cases. In contrast, directed evolution has been applied to alter the stabilities and enzymatic activities of a number of different proteins. Given an

appropriate screening or selection system, directed evolution is generally applicable, because it does not require information on the structure of the protein that is to be stabilized. The directed evolution approach is highly instructive, because it is unbiased and often provides unexpected solutions to a given problem. Because wild-type and altered proteins differ only by few amino acids, the analysis of the mechanisms of stabilization and activation is considerably simplified, compared with the analysis of more divergent pairs of homologous mesophilic and (hyper)thermophilic proteins.

VI. PROTEIN THERMOSTABILITY VERSUS PROTEIN FLEXIBILITY

Hyperthermophilic enzymes are often barely active at room temperature, but are as active as their mesophilic counterparts at the corresponding physiological temperatures (Jaenicke, 1991). Because the active site residues are conserved between homologous mesophilic and hyperthermophilic enzymes, these differences in catalytic properties must be caused by substitutions elsewhere in the molecule. For each member of a homologous enzyme series, a certain degree of conformational fluctuations is required for optimal catalysis. Therefore, it has been postulated that the low activity of thermostable enzymes at low temperatures is due to a high rigidity, which is relieved at the elevated temperatures at which hyperthermophilic enzymes work *in vivo* (Jaenicke, 2000). In accordance with this hypothesis, the activity of sTrpC was increased by increasing the flexibility of an active site loop (Merz et al., 2000, cf. Section V.B) and various T4 lysozyme variants with increased stabilities show an increased rigidity, as reflected in decreased crystallographic thermal factors (Wray et al., 1999), or in a reduced catalytic activity (Shoichet et al., 1995).

Isopropylmalate dehydrogenase (IPMDH) was used as a model to investigate the correlation between stability, flexibility, and catalytic activity for a pair of mesophilic and thermophilic enzymes (Závodszky et al., 1998). IPMDH from *E. coli* (eIPMDH) and *T. thermophilus* (ttIPMDH) are homodimeric enzymes with an rms deviation between equivalent Cα-atoms of only 1.58 Å. The catalytically essential amino acid residues are fully conserved between the two enzymes. However, the temperature for optimum specific activity is about 50°C for eIPMDH but about 74°C for ttIPMDH, which is only marginally active at room temperature. eIPMDH and ttIPMDH were compared with respect to the thermodynamic stabilities ΔG_{FU} of the entire protein (Section II.A), which was termed "macrostabilty" here, and with respect to the thermodynamic stabilities of protein substructures, which were termed "microstability". Whereas the macrostability maintains the integrity of the native state, the microstability determines the flexibiliy or local conformational dynamics of a protein, which is presumably responsible for optimum function. The measurement of the macrostabilities of eIPMDH and ttIPMDH by differential scanning calorimetry and circular dichroism yielded T_m values of about 74 and 90°C, respectively (Figure 11). The microstabilities are reflected by global hydrogen exchange kinetics, which were measured using Fourier-transformed infrared (FT-IR) spectroscopy. Reversible, conformational fluctuations expose buried segments of the polypeptide chain to the solvent. When the protein is dissolved in D_2O, hydrogen-deuterium (H/D) exchange of peptide amide protons occurs during such exposure (Equation 13). The H/D exchange analysis was performed under EX_2 conditions, where fluctuations exposing buried hydrogens are fast compared to the exchange rate of the solvent-exposed peptide groups. When com-

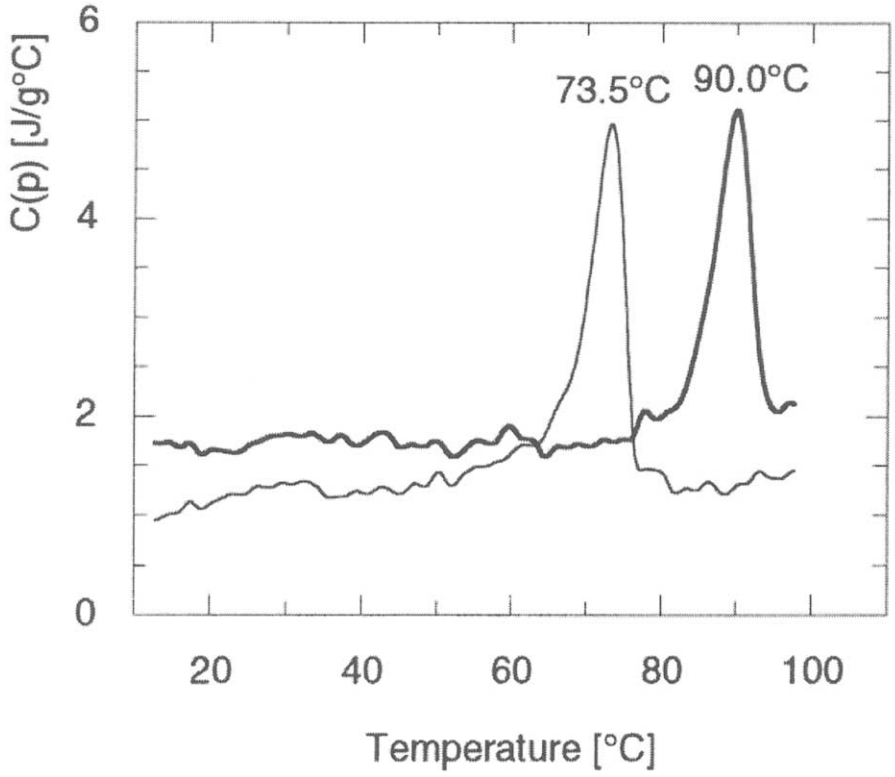
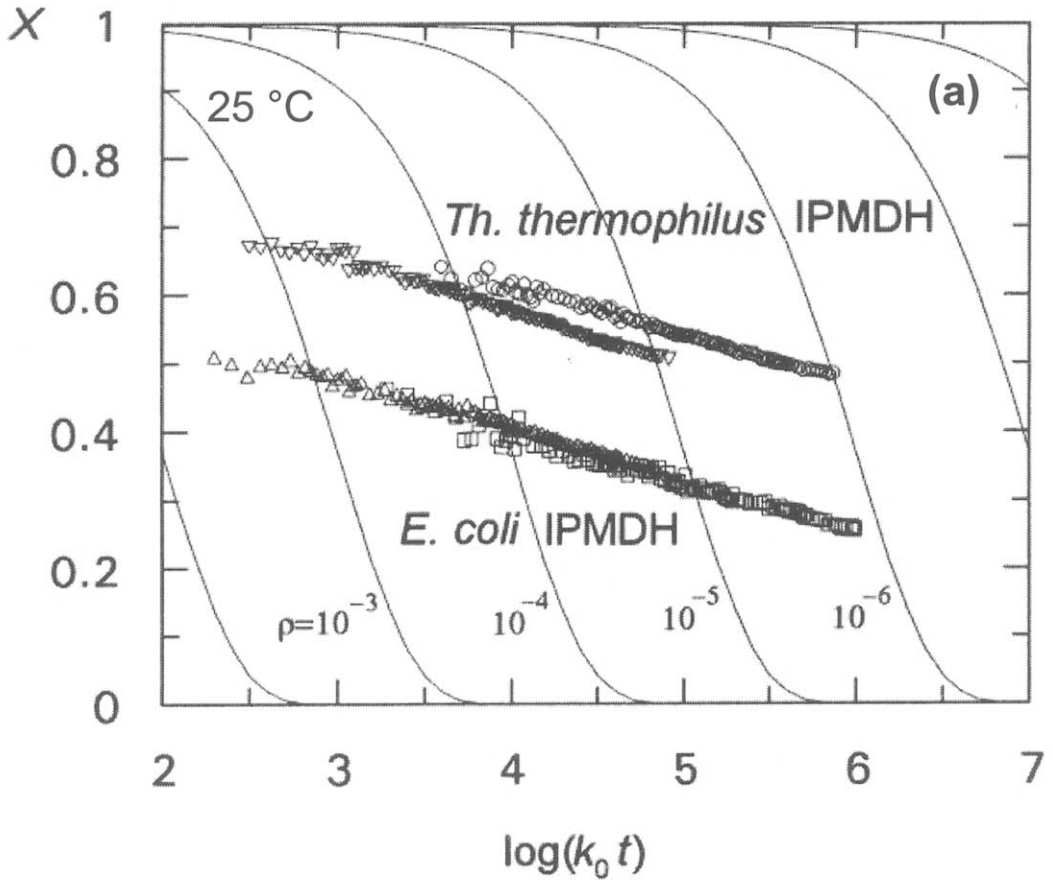


FIGURE 11. Isopropylmalate dehydrogenase from *Escherichia coli* (eIPMDH; thin line; T_m =73.5°C) is less thermostable than its homologue from *Thermus thermophilus* (ttIPMDH; thick line; T_m = 90.0°C). The shown calorimetric melting profiles were measured with a heating rate of 1°C/min in 20 m*M* potassium phosphate, pH 7.6. KCl concentrations were set to 0.3 *M* and 1.0 *M*, respectively. (Reprinted with permission from Závodszky, P., Kardos, J., Svingor, and Petsko, G. A. *Proc. Natl. Acad. Sci. USA* **95:** 7406–7411. Copyright (1998) National Academy of Sciences, U.S.A.)

pared at 25°C, the hydrogen protection factors of ttIPMDH are significantly higher compared to eIPMDH, that is, ttIPMDH is significantly less flexible according to this criterion (Figure 12a). In contrast, at temperatures near the activity optima (70°C for ttIPMDH and 48°C for eIPMDH), the protection factors are similar (Figure 12b). These findings support the hypothesis that the low specific activity of ttIPMDH at room temperature is due to its conformational rigidity.

Rubredoxin from *Pyrococcus furiosus* (pRdx) is the most thermostable protein characterized to date, with an extrapolated T_m value of almost 200°C and an estimated global unfolding rate of 10^{-6} s^{-1} at 100°C (Hiller et al., 1997; Cavagnero et al., 1998; cf. Section II.C). Hernandez et al. (2000) monitored the magnetization transfer rate of ^1H spins migrating between the ^1H_2O and $^1H^N$ frequencies in order to quantify the exchange of individual amide hydrogens. The Zn^{2+} form of ^{15}N-labeled pRdx was investigated at pH-values from 7.17 to 12.51 and at 3°C, 28°C, and 53°C. The EX_2 mechanism was verified for the majority of the protein amides. Surprisingly, conformational fluctuations that led to hydrogen bond rupture and access of solvent occur in the millisecond time scale throughout the entire protein. At 28°C, the local opening rates k_{op} (cf. Equation 13) are about 10^3 s^{-1} on average and about 70 s^{-1} for the slowest exchanging hydrogens. The corresponding



A

FIGURE 12. Flexibilities of eIPMDH and ttIPMDH at various temperatures. (a) At 25°C, the structure of eIPMDH is more flexible than that of ttIPMDH. (b) The structures of eIPMDH at 48°C and of ttIPMDH at 70°C are comparably flexible. H/D exchange data, summarized in the form of relaxation spectra for both eIPMDH (△, pD = 7.15; □, pD = 8.15) and ttIPMDH (▽, pD = 7.15; ○, pD = 8.15). *X* is the fraction of unexchanged peptide hydrogens, *t* is the time, and k_0 is the chemical exchange rate constant. The solid lines represent the exchange rate curves for hypothetical polypeptides characterized by the indicated ρ values, which are a measure for the probability of solvent exposure of the peptide groups. (Reprinted with permission from Závodszky, P., Kardos, J., Svingor, and Petsko, G. A. *Proc. Natl. Acad. Sci. USA* **95:** 7406–7411. Copyright (1998) National Academy of Sciences, U.S.A.)

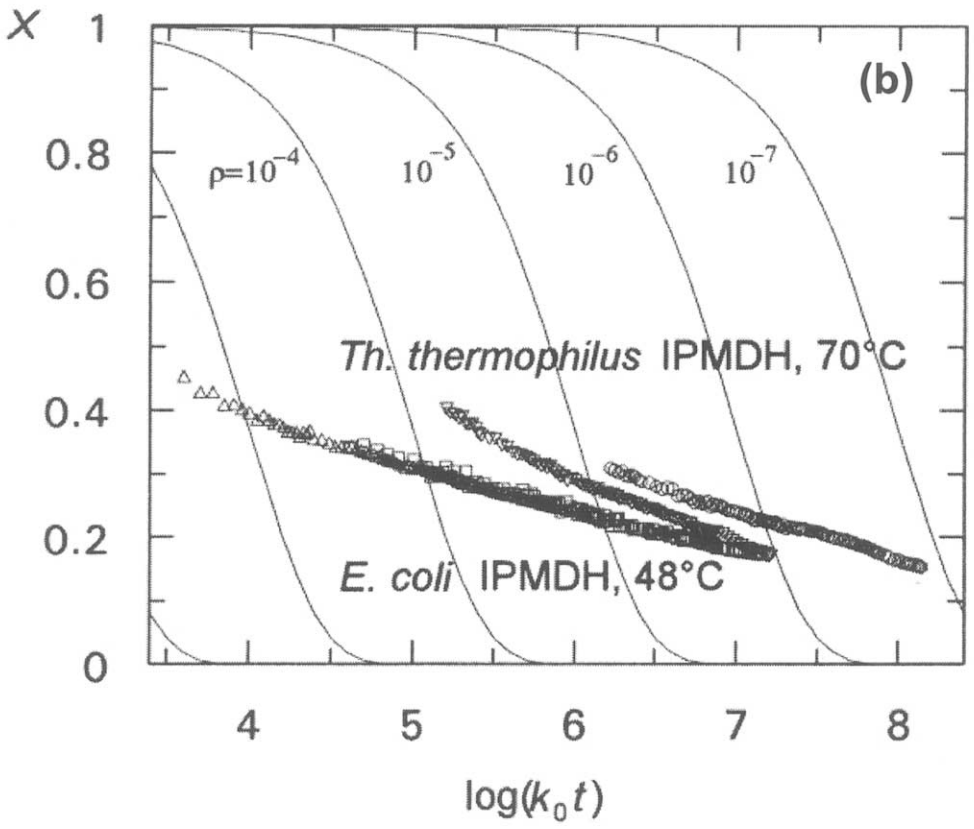


FIGURE 12 B

amide protection factors for pRdx are typical of those observed for a large number of mesophilic proteins. As a consequence, the microstability of this hyperthermophilic protein is indistinguishable from that of typical mesophilic proteins. Also, the local opening rates in pRdx of about 10^3 s^{-1} at 28°C are in marked contrast to the global unfolding rate of 10^{-6} s^{-1} at 100°C, that is, conformational fluctuations required for amide exchanged monitored herein and conformational fluctuations required for global unfolding must be very weakly coupled.

These results are in striking contrast to what has been discussed above for eIPMDH and ttIPMDH (Závodszky et al., 1998). The reasons for this discrepancy are not clear, but the damaging of pRdx or drastic alterations of the dynamics at the applied extremely high pH values have been discussed (Jaenicke, 2000). On the other hand, there are extremely stable enzymes from hyperthermophiles that are catalytically highly active at low temperatures (Sterner et al., 1996), suggesting that flexibility at the active site and overall stability of the protein are not necessarily mutually exclusive. Along these lines, it has been shown for a number of enzymes, either by site-directed mutagenesis or by directed evolution, that high catalytic activity at low temperatures and high thermostability can be combined in the same enzyme (van den Burg et al., 1998; Giver et al., 1998; Williams et al., 1999; Song and Rhee, 2000; Lehmann et al., 2000).

The rarely found combination of high thermostability and high catalytic activity at low temperature in native enzymes is probably due to lacking evolutionary constraints (Miyazaki et al., 2000). Enzymes in psychrophilic or mesophilic organisms are under no selective pressure to be stable at elevated temperature. Similarly, the activities of hyperthermophilic enzymes at low temperatures normally need not be high. Otherwise, hyperthermophilic enzymes would be much more efficient catalysts than their mesophilic homologues at the corresponding physiological temperatures, due to the inevitable acceleration of chemical reactions with increasing temperature.

VII. PROTEIN THERMOSTABILITY *IN VIVO*

Considering the high energetic cost of protein synthesis, keeping proteins in the native state is favorable compared with new synthesis. Exceptions are proteins that, due to the needs of metabolic regulation, have to be turned over rapidly. The need to stabilize proteins obviously becomes more important the closer the temperature approaches (1) the upper growth limit of an organism and (2) the temperature range where thermal decomposition of amino acids and the protein backbone become significant. During evolution of cellular life, mechanisms have been developed to enhance the *in vivo* stability of proteins beyond their intrinsic stability by additional factors such as sophisticated molecular chaperone machines, protein repair enzymes and high concentrations of small stabilizing solutes.

A. Chaperones Aid in Folding and Refolding Under Heat Stress

Chaperones play essential roles in prokaryotic as well as eukaryotic cells. They are involved in the proper folding of newly synthesized proteins and of polypeptides translocated through membranes, they can prevent misfolding and aid in the refolding of denatured proteins. Some chaperones actively unfold misfolded proteins, thereby routing them either to a correct fold or to pro-

teolytic degradation (Horwich et al., 1999). Chaperones belong to the class of heat-shock proteins (Hsp), because their synthesis in the cell is often induced by a temperature upshift. Typically, the action of cytoplasmic chaperones is dependent on the hydrolysis of ATP. Chaperones are named according to their subunit masses in kilodaltons (Hsp100, Hsp90, Hsp70, Hsp60, Hsp40, and small Hsps). Most information is available for the Hsp70 and the Hsp60 chaperone families. The Hsp60 chaperons, which are also called chaperonins, form large double-ring assemblies with a central cavity. Folding that is assisted by either of the major chaperone systems, Hsp70 or Hsp60, may require multiple rounds of polypeptide binding and release (Bukau and Horwich, 1998; Netzer and Hartl, 1998; Macario et al., 1999).

It appears that chaperones play a crucial role near the upper temperature border of hyperthermophilic life. *Pyrodictium occultum* has optimum and maximum growth temperatures of 105 and 110°C, respectively. At 108°C, heat-induced protein synthesis results in the overproduction of a 16-subunit chaperone to approximately 80% of the soluble cellular protein (Phipps et al., 1993). The extremely high concentration of this "thermosome" indicates that the organism must spend a substantial proportion of its metabolic energy for the folding and refolding of its proteins. Heat-induced synthesis of chaperonins has also been reported for *Archaeoglobus fulgidus* and *Sulfolobus shibatae* (Emmerhoff et al., 1998; Kagawa et al., 1995), but was not observed in *Methanopyrus kandleri* (Minuth et al., 1999). In the following sections, the state of our knowledge about chaperone distribution, structure, and function is summarized, with an emphasis on chaperones found in extreme thermophiles. A more general overview is given only for the Hsp70 and Hsp60 chaperones, but reference to other chaperones is made where appropriate.

1. Hsp70

Hsp70 homologues are found in all three domains of life, the *Eucarya* (in the cytosol, mitochondria, chloroplasts, and endoplasmic reticulum), the *Bacteria* and the *Archaea*, although this chaperone class is not present in all archaea (see below). In bacteria, the alternate name DnaK is often used instead of Hsp70. Hsp70 chaperones have a multitude of functions, including the binding and assistance in folding of nascent polypeptide chains, prevention of aggregation, refolding of misfolded proteins, oligomer disassembly, and assistance in translocation of peptide chains through organelle membranes (Bukau and Horwich, 1998; Deuerling et al., 1999; Hesterkamp and Bukau, 1998; Mogk et al, 1999; Schröder et al., 1993). Hsp70 chaperones have a detergent-like function, separating hydrophobic peptide regions, either as they appear on nascent chains during translation, or after aggregation. Nucleotide binding and hydrolysis, which is modulated by co-chaperones, is crucial for peptide binding and release (Bukau and Horwich, 1998). The concerted function of Hsp70 (DnaK) with the co-chaperone Hsp40 (DnaJ) and the nucleotide-exchange factor GrpE is well documented for the bacterial system (Netzer and Hartl, 1998; Rüdiger et al., 1997), but little information is available for the archaea. At 30 to 37°C, DnaK from *E.coli* can be functionally replaced by the ribosome-associated chaperone and peptidyl-prolyl *cis/*trans isomerase trigger factor (Deuerling et al., 1999; Teter et al., 1999). Under stress conditions, however, DnaK is required for cell survival and for the refolding of certain denatured proteins (Hesterkamp and Bukau, 1998; Schröder et al., 1993).

Some archaea, such as *Methanobacterium thermoautotrophicum* or *Thermoplasma acidophilum*, possess Hsp70 (DnaK),

which, like in the bacteria, are encoded by genes clustered together with Hsp40(DnaJ) and GrpE gene homologues. Other archaea, including all of the most thermophilic representatives of this domain (genera *Sulfolobus, Methanococcus, Meth-anothermus, Archaeoglobus, Desulfur-ococcus, Thermococcus, Methanopyrus, Pyrococcus, Pyrobaculum, Aeropyrum*) have no Hsp70-homologue encoding genes in their genomes (Gribaldo et al., 1999; Macario et al., 1999). A different and structurally unrelated chaperone may fulfill Hsp70-equivalent functions in these species (Leroux et al., 1999; Siegert et al., 2000). In contrast, thermophilic and hyperthermo-philic bacteria, including the genera *Thermus, Thermotoga*, and *Aquifex*, contain Hsp70 homologues (Macario et al., 1999; genome sequence data). The question whether archaea lacking Hsp70 contain a functional equivalent is especially important when considering the cooperation of Hsp70 with other chaperone systems, for example, with Hsp60 (GroEL) as demonstrated for *E. coli*, or with Hsp100/Clp as demonstrated for *E. coli* and yeast (see below).

2. Hsp60

The Hsp60 (chaperonin) complexes of bacteria are composed of two heptameric rings of GroEL subunits, which form a functional complex with one heptameric ring of the co-chaperone GroES. The GroEL chaperonin resembles a hollow cylinder, the central cavity of which can accommodate a single polypeptide with a maximum size of about 55 kDa (Netzer and Hartl, 1998). Binding of the substrate protein takes place via hydrophobic patches on the inside of the cylinder. Binding of ATP and sealing of the entrance hole results in a dramatic conformational change, releasing the polypeptide into an isolated hydrophilic compartment ('Anfinsen cage' or folding cage). Here, folding can take place in an environment that shields the nascent polypeptide chain from aggregation with other macromolecules in the cytosol. GroEL/GroES may cooperate with the Hsp70(DnaK)/Hsp40(DnaJ)/GrpE system to assure correct folding, whereby the Hsp70 system maintains the polypeptide in a folding-competent state and the final steps are facilitated by GroEL/GroES. No comparable interactions between the corresponding Hsp70 and Hsp60 systems of the eukaryotic cytosol were found. Interaction of Hsp60 and Hsp70 in those archaeal species containing Hsp70 has not been explored yet (Macario et al., 1999; Netzer and Hartl, 1998).

The Hsp60 complexes of thermophilic and hyperthermophilic archaea are called thermosomes. They were found in a number of species such as *Thermoplasma acidophilum, Pyrodictium occultum, Pyrodictium brockii, Pyrococcus* sp., *Methanococcus jannashii, Methanopyrus kandleri, Sulfolobus solfataricus*, and *Thermococcus* sp. (Andrä et al., 1998; Gutsche et al., 1999; Joachimiak et al., 1997; Kowalski et al., 1998; Minuth et al., 1998, 1999; Quaite-Randall et al., 1995 ; Waldmann et al., 1995; Yan et al., 1997; Yoshida et al., 1997). The cloning of homologous *hsp*60 genes and their detection in whole genome sequencing projects indicate that similar chaperonins also exist in the following thermophilic archaea: *Archaeoglobus fulgidus, Methanococcus jannaschii, Methanococcus thermolithotrophicus, Methanobacterium thermoautotrophicum, Desulfurococcus* sp., *Pyrobaculum aerophilum, Aeropyrum pernix, Pyrococcus furiosus, Pyrococcus kodakaraensis, Pyrococcus abyssi, Sulfolobus acidocaldarius*, and *S. shibatae* (Archibald et al., 1999 ; Macario et al.,

1999). Archaeal thermosomes have comparable functional properties as their bacterial Hsp60 counterparts. They consist of multiple subunits that form double-stacked rings, display ATPase activity and bind unspecifically to nonnative proteins *in vitro* (Andrä et al., 1998; Furutani et al., 1998; Minuth et al., 1998, 1999; Yan et al., 1997). However, striking structural differences exist between the thermosomes and bacterial Hsp60 complexes such as GroEL from *E. coli*. The layers of the archaeal Hsp60s contain eight or nine subunits that belong to 2 to 3 different but related types (Archibald et al., 1999). Thus, the archaeal thermosome structures are similar to chaperonin complexes found in the cytosol of eucaryotes, which are called "cytosolic chaperonin containing TCP-1 (CCT) or TCP-1 ring complex" (TRiC) (TCP-1 stands for 'tailless complex polypeptide 1'). TRiC consists of eight different subunits in contrast to the homo-oligomeric heptamer rings of bacterial GroEL. Also, thermosome action does not depend on a GroES-like co-chaperonin ring, which forms a lid that alternatively seals either the top or the bottom side of the GroEL barrel. Instead, in the *Thermoplasma* thermosome, access to the central cavity appears to be controlled by a built-in lid consisting of protrusions of the apical chaperonin domains (Klumpp et al., 1997; Ditzel et al., 1998). Also in this respect the archaeal Hsp60 chaperonins more closely resemble the chaperonin of the eukaryotic cytosol. The type of chaperonin present in the eucarya and archaea is called group II chaperonin to differentiate it from the GroEL/GroES-like group I chaperonin of the bacteria.

Hyperthermophilic as well as mesophilic archaea contain the group II chaperonin complex, suggesting that it is not a specific adaptation to high temperatures. Along these lines, the genomes of the hyperthermophilic bacteria *Thermotoga maritima* and *Aquifex aeolicus* both contain gene homologues of group I chaperonin systems, just like the nonthermophilic bacteria (Deckert et al., 1998; Nelson et al., 1999). Group II chaperonins interact with the recently discovered 'prefoldin' (GimC), an ATP-independent chaperone that may fulfill Hsp70-equivalent functions in the archaea. Gene homologues for this factor are present in all archaea but are missing in the bacteria (Siegers et al., 1999; Leroux et al., 1999; Leroux and Hartl, 2000). GimC exists as a hexameric complex and, in contrast to GroES, has its own chaperone activity in that it can bind to and stabilize nonnative proteins (Siegers et al., 1999). Apparently, flexible tentacle-like protrusions, which expose hydrophobic patches at their tips, are involved in the binding of nonnative peptide segments by GimC (Siegert et al., 2000). The precise role of GimC during its interaction with group II chaperonins in the course of protein folding remains to be defined.

Recent work also points to strong differences between group I and group II chaperonins with respect to their allosteric properties and their nucleotide binding/hydrolysis reaction cycle. In contrast to group I chaperonins, no cooperativity of nucleotide binding to the *Thermoplasma* thermosome was observed. This thermosome, which consists of two different subunit types, apparently has two classes of ATP-binding sites with different binding affinity. ATP hydrolysis, rather than ATP binding as observed with GroEL/GroES, triggers the closure of the thermosome folding cage (Gutsche et al., 2000a,b).

Relatively little is known about the functional mechanisms of thermosomes from hyperthermophiles. The *Pyrodictium occultum* thermosome is composed of two octameric rings, each consisting of alternating α and β subunits. *In vitro* studies showed that homo- and hetero-oligomeric forms consisting of the α and/or β subunits can interact

with non-native protein substrates and prevent their irreversible aggregation. However, release of the bound proteins from the chaperonin-substrate complexes was not possible, and binding could not be modulated by Mg-ATP under *in vitro* assay conditions (Minuth et al., 1998). Similarly, the homohexadecameric thermosome of *Methanopyrus kandleri* displayed chaperone-like activity *in vitro,* but led to the production of dead end complexes with nonnative substrates (Minuth et al., 1999). These experiments, however, were carried out with heterologous, mostly mesophilic substrate proteins, and far below the optimum growth temperatures of the thermosome-producing organisms. Thus, further experiments with physiological substrates at high temperatures are necessary. However, the available protein substrates from extreme thermophiles often appear to be unsuited, either because they are too stable to be unfolded, or because their high refolding rates pose problems for kinetic analyses (Minuth et al., 1999).

3. Prevention of Protein Aggregation Under Thermal Stress

Molecular chaperones can prevent the aggregation of proteins that unfold after exposure of the cell to high temperature. This function of chaperones explains why the expression of many chaperone genes is induced by a temperature up-shift. Clearly, not all proteins require protection against aggregation, as some resist thermal unfolding at temperatures well above the maxiumum growth temperature of the producing organism (Pfeil et al., 1997; Wassenberg et al., 1997, 2000). However, even if only a few essential proteins are prone to aggregation under heat stress, the cell must provide an efficient antiaggregation/refolding apparatus in order to survive a temperature up-shift. The situation probably gets more severe as the temperature approaches the upper limit for growth of an organism.

It has been estimated that about 10 to 15% of the *E. coli* proteins were denatured and aggregated during a 15-min incubation of cell extracts at 45°C, a temperature that still promotes growth of this organism (Goloubinoff et al., 1999; Mogk et al., 1999). The aggregated protein fraction consisted of about 250 protein species. Aggregation of most proteins was at least partially prevented by addition of DnaK to the extracts before heat treatment. Also, the exposure of a Δ*dnaK::cat* strain to a mild heat shock (60 min at 42°C) resulted in strong protein aggregation *in vivo*, as detected by two-dimensional gel electrophoresis. In contrast, in DnaK-containing wild-type cells no aggregation was detected (Mogk et al., 1999).

Similar experiments carried out with strains containing mutations in other chaperone genes (knockout mutations in *clpB, ibpAB, htpG*; conditional mutations in *groEL)* indicated that the DnaK system is the most effective chaperone system in preventing protein aggregation in *E. coli* cells. Mass spectrometric analysis of a number of aggregated proteins from the heat-shocked Δ*dnaK::cat* strain showed that 80% of the large (≥90 kDa) proteins were thermolabile, aggregation-prone, and substrates for DnaK. In contrast, only 18% of the small (≤30 kDa) proteins showed a heat-induced tendency to aggregate (Mogk et al., 1999). The fact that most large proteins tend to denature and aggregate even under mild heat shock conditions underscores the importance of an efficient chaperone system for cell survival.

4. Solubilization and Refolding of Aggregated Proteins

Protein unfolding is an important function of chaperones and chaperone networks. On the one hand, unfolding is important to

feed misfolded and denatured proteins into the cellular protein degradation system (Horwich et al., 1999; Weber-Ban et al., 1999). On the other hand, protein unfolding induced by chaperones could be beneficial for organisms when growing near their upper temperature limit, because it offers aggregated proteins a new chance for proper folding. For the bacterial GroEL chaperonin system it has been suggested that unfolding could occur during the release of the multivalently bound polypeptide substrate via twisting movements of GroEL domains brought about by binding of ATP-GroES (Bukau and Horwich, 1998; Lorimer, 1997; Roseman et al., 1996; Shtilerman et al., 1999). The unfolding function may be of general relevance for GroEL-aided protein folding, but its possible role in the management of heat-induced damage is unclear, because GroEL/GroES-mediated reactivation of substrate enzymes *in vitro* depends on a large molar excess of chaperones (Veinger et al., 1998; Ziemienowicz et al., 1993). Thermophilic archaeal group II chaperonins apparently also undergo dramatic domain rotations during their functional cycle (Gutsche et al., 2000c; Schoen et al., 2000), which makes it plausible that they can also mediate unfolding reactions. However, in this case the substrate binding patches on the inside chaperonin wall have not yet been mapped.

Recent work has demonstrated the remarkable unfoldase activity of members of the Hsp100/Clp family of chaperones in yeast and in *E. coli*, and their cooperation with the Hsp70 chaperone system to solubilize and subsequently refold previously aggregated proteins (Glover and Lindquist, 1998; Goloubinoff et al., 1999; Mogk et al., 1999; Weber-Ban et al., 1999). Protein aggregates are first bound by ClpB, and an ATP-triggered structural change leads to the presentation of hydrophobic aggregate regions that are subsequently solubilized by the DnaK system (Figure 13). The broad substrate spectrum of this system was demonstrated by the efficient solubilization (55%) of a number of aggregated model enzymes, such as mitochondrial malate dehydrogenase, firefly luciferase, glucose-6-phosphate dehydrogenase, and α-glucosidase. In the absence of ClpB, the DnaK system was able to solubilze only 5 to 6% of the aggregated proteins (Goloubinoff et al., 1999). Synergistic cooperation was also demonstrated for ClpB and the DnaK system of the extremely thermophilic bacterium *Thermus thermophilus*. *In vitro*, the DnaK system (DnaK, DnaJ, GrpE and ATP) suppressed heat-induced aggregation of substrate proteins, and efficient reactivation was observed in the presence of the *Thermus* ClpB homologue (Motohashi et al., 1999). Furthermore, the addition of *Thermus* chaperonin improved the yield of reactivation (Watanabe et al., 2000). Interestingly, the *T. thermophilus* ClpB homologue is encoded within the *dnaK* region, in a gene cluster that also contains the genes for DnaK, DnaJ, GrpE, and DafA. DafA (DnaK-DnaJ assembly factor) is part of a peculiar $DnaK_3$-$DnaJ_3$-$DafA_3$ trigonal ring complex found in this thermophilic organism (Motohashi et al., 1999). Apparently, DafA stabilizes the inactive ternary complex until it is replaced by a substrate protein, yielding a DnaK-substrate-DnaJ complex, which is the active chaperone species (Klostermeier et al., 1999). At present it is unclear whether similar ClpB-Hsp70 networks exist in other extreme thermophiles. The genomes of the hyperthermophilic bacteria *Thermotoga maritima* and *Aquifex aeolicus* have the coding capacity for homologues of the Hsp70 system as well as for ClpB, but there are no experimental data on their function and possible cooperation. To our knowledge no ClpB ho-

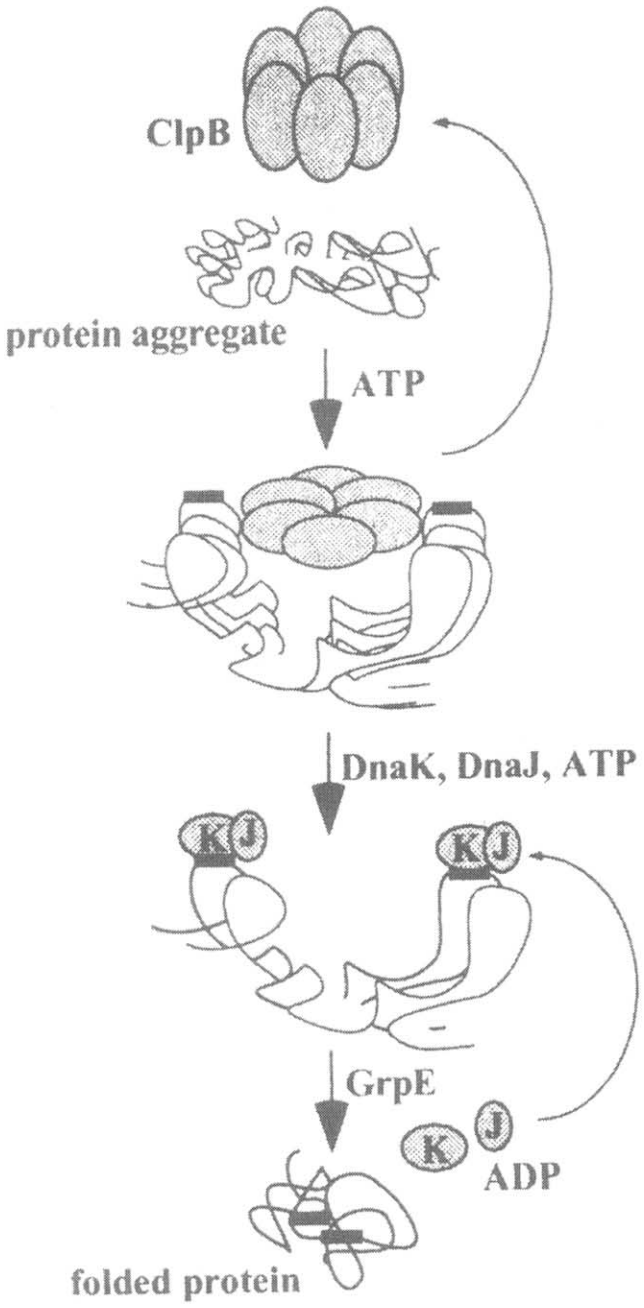


FIGURE 13. Model for the cooperative and sequential action of the chaperone ClpB with the chaperone system DnaJ-DnaK-GrpE of *E. coli* for the catalytic solubilization and refolding of protein aggregates. (Reprinted with permission from Goloubinoff, P., Mogk, A., Peres Ben Zvi, A., Tomoyasu, T., and Bukau, B. *Proc. Natl. Acad. Sci. USA* **96:**13732–13737. Copyright (1999) National Academy of Sciences, U.S.A.)

mologues have been found in the archaea so far.

5. Small Heat Shock Proteins

An interesting but not well-studied group of chaperones are the α-crystallin-like small heat shock proteins (sHsp), exemplified in the *Bacteria* and the *Archaea* by *E. coli* IbpA and IbpB, and *Methanococcus jannaschii* Hsp16.5, respectively. Structural information is available for some sHsps (Haley et al., 2000; Kim et al., 1998a). *M. jannaschii* sHsp is organized as a 24-subunit homo-multimeric hollow sphere with small 'windows'. Proteins from *E. coli* extracts containing the *M. jannaschii* sHsp remained soluble up to 100°C, while proteins in sHsp-free extracts precipitated at 60°C. Also, purified proteins were protected from aggregation *in vitro* (Kim et al., 1998a,b). *E. coli* IbpB prevented the irreversible aggregation of heat-denatured malate dehydrogenase and chemically denatured lactate dehydrogenase. This sHsp aids in refolding substrate proteins by delivering them to the DnaK/DnaJ/GrpE chaperone system (Veinger et al., 1998). Details of the functional mechanism of sHsps remain to be elucidated.

6. Summary

Our knowledge about chaperone action is rapidly advancing, in particular in the exciting area of how different chaperone systems cooperate synergistically with each other to fold, unfold, disaggregate, solubilize, and reactivate proteins. The fact that the production of many chaperones is induced after heat shock clearly demonstrates that they are important to cope with this stressor. It is plausible that they become more and more important as an organism grows at its upper temperature limit, where at least part of its proteins undergo heat denaturation at a significant rate. Proteins and protein complexes with chaperone-like properties have been found in various (hyper)thermophilic microorganisms, and significant progress made in recent years sheds light on their structural and functional characteristics. However, only little information is available about their role *in vivo* at temperatures near the upper temperature border of life.

B. Protein Repair Enzymes

In a sense chaperones can be regarded as repair enzymes, because they can catalyze the refolding of misfolded polypeptides (see above). A different category of protein repair enzymes can counteract detrimental covalent modifications of polypeptides, which take place at considerable rates at the high physiological temperatures of thermophiles and hyperthermophiles. Such modifications include deamidation of the amide side chain of Asn and Gln residues, succinimide formation at Glu and Asp residues, and oxidation of His, Met, Cys, Trp, and Tyr residues (Daniel et al., 1996; Geiger and Clarke, 1987; Stephenson and Clarke, 1989). A particular problem is the deamidation of the side chains of Asn and Gln residues, and, as a consequence, the fraction of these amino acids is significantly decreased in hyperthermophilic proteins (Section III.A). The deamidation of Asn residues is initiated by the formation of a succinimide ring. This ring is hydrolyzed in aqueous solution to yield aspartyl as well as isoaspartyl residues, which disturb the protein backbone. The reaction is pH dependent, and the nature of neighboring amino acid residues can affect the rate of deamidation (Wright, 1991). In mesophilic, actively grow-

ing cells deamidation rates are normally slow and thus can be compensated by normal protein turnover, but in nongrowing cells such as spores, stationary phase cells, and at elevated temperature deamidated proteins can accumulate.

The susceptibility of a protein to deamidation and other degradative reactions is linked to its conformational integrity. For example, when compared with the native enzyme, denatured glyceraldehyde-3-phosphate dehydrogenase from *Pyrococcus woesei* displayed an increased rate of deamidation at 100°C, just like a labilized mutant form of glyceraldehyde-3-phosphate dehydrogenase from *Methanothermus fervidus* (Hensel and Jakob, 1994). It appears that most of the vulnerable residues are more or less well protected in the native polypeptide and that irreversible deamidation, which can cause protein inactivation (Brennan et al., 1994; Hayes and Setlow, 1997), occurs preferably in the unfolded state of a protein (Section II.B and Equation 12). The rate of deamidation can become very high at physiological temperatures of hyperthermophiles and pose a serious problem for protein stability *in vivo*.

The enzyme L-isoaspartyl *O*-methyltransferase counteracts the accumulation of abnormal isoaspartyl residues formed by the deamidation of proteins. It catalyzes the transfer of a methyl group from *S*-adenosylmethionine to the α-carboxyl group of L-isoaspartyl residues to initiate the 'repair' of these residues to aspartyl residues (Brennan et al., 1994; Johnson et al., 1987). The reactions following the enzymatic L-isoaspartyl methylation are nonenzymatic. Spontaneous demethylation occurs to yield a cyclic aspartyl imide intermediate, which subsequently becomes hydrolyzed to a mixture of L-isoaspartyl and L-aspartyl residues (Johnson et al., 1987). These reactions convert the abnormal isopeptide bond to a normal peptide bond. However, the asparagine side chain amide is not restored. L-isoaspartyl methyltransferase is a conserved protein that has been identified in animals, plants, and bacteria. *E. coli* mutants lacking this enzyme are affected in stationary-phase survival and heat-shock resistance (Li and Clarke, 1992). Homologous genes were also identified in hyperthermophilic bacteria and archaea (Ichikawa and Clarke, 1998). The corresponding gene from *Thermotoga maritima* was expressed heterologously, and the gene product was purified and characterized. The enzyme is extremely resistant against thermoinactivation, as it retains its full catalytic activity after an incubation at 100°C for 60 min. Remarkably, the enzyme displayed a nearly 20-fold higher specific activity at the physiological temperature of 85°C than mesophilic enzymes at 37°C (Ichikawa and Clarke, 1998). L-isoaspartyl methyltransferase might thus be important for the avoidance of protein inactivation during growth of *Thermotoga maritima* at high temperature.

C. Compatible Solutes

Low-molecular-mass solutes in the aqueous environment of proteins can have various effects on protein solubility and stability. At low concentrations, salts can increase protein solubility ('salting-in'), whereas at high concentrations salts can lead to protein precipitation ('salting-out'). Also, some solutes (for example, urea or guanidinium chloride) destabilize proteins, whereas others (for example, sucrose or glycerol) stabilize them (Timasheff, 1993, 1995). All these effects require a high concentration (usually >1 *M*) of the added compound. Therefore, the interactions involved must be nonspecific and weak. Water plays an important role here, because the effect of a particular

compound depends on the differential affinity of the protein for water and the agent (Timasheff, 1993). The precipitating and stabilizing compounds are preferentially excluded from the surface of the protein, that is, the protein has a higher affinity for water than for these agents. As a consequence, proteins are preferentially hydrated, which favors the native state and makes unfolding more unfavorable. In contrast, destabilizing agents bind more strongly than water to proteins. Stabilizing compounds include sugars and polyols (sucrose, trehalose, glycerol, mannitol, sorbitol), amino acids (proline, glycine), methyl amines (sarcosine, trimethylamine-*N*-oxide, glycine betaine), tetrahydropyrimidine derivatives (ectoins), some salting-out salts (Timasheff, 1995; Knapp et al., 1999), as well as a number of compounds that have been found specifically in thermophiles and hyperthermophiles (Figure 14).

Many organisms accumulate high concentrations of organic solutes in response to stress conditions such as high osmolarity, desiccation, freezing, or high temperature. These solutes are called compatible solutes, because they do not compromise vital cellular functions even when present at high concentrations (DaCosta et al., 1998). Early studies dealing with the properties of enzymes from hyperthermophiles demonstrated a surprisingly low *in vitro* resistance against thermoinactivation for some of these proteins (Thomm et al., 1986; Fabry and Hensel, 1987). These observations suggested that extrinsic stabilizing compounds could be involved in *in vivo* protein stabilization and thermoadaptation of (hyper)thermophiles and led to the search for protective compatible solutes in these organisms. In the following, the knowledge about compatible solutes that were recently discovered in thermophiles and hyperthermophiles are summarized. Experiments to prove that these new compatible solutes are specific adaptations to life at extreme temperatures need to demonstrate that (1) they are important or even essential for growth at high temperatures, (2) they are produced or overproduced as a response to a temperature up-shift, and/or (3) they perform better than other known compounds from mesophiles in protein stabilization. Unfortunately, mutants in the biosynthetic pathways for the compounds are not available, making it impossible to assess if they are essential. On the other hand, an increase of the intracellular concentration of several of these compounds in response to a temperature increase has been demonstrated, and some of them protected selected proteins against thermoinactivation *in vitro* (see below). Although mainly the thermophiles and hyperthermophiles are discussed, it should be kept in mind that also compounds produced by mesophiles (for example, trehalose or ectoines) can accumulate in response to heat stress, can increase the thermotolerance of the producing microorganism, and can exert stabilizing effects on cellular components in *in vitro* experiments (Argüelles, 2000; Knapp et al., 1999; Lippert and Galinski, 1992; Malin and Lapidot, 1996).

1. Cyclic 2,3-Diphosphoglycerate

Cyclic 2,3-diphosphoglycerate (cDPG; Figure 14a) was first detected in the moderately thermophlic bacterium *Methanobacterium thermoautotrophicum* (Kanodia and Roberts, 1983; Seely and Fahrney, 1983). In some hyperthermophilic methanogens, including (growth temperatures in parenthesis) *Methanothermus fervidus* (84°C), *Methanothermus sociabilis* (88°C), and *Methanopyrus kandleri* (98°C), cDPG accumulates to concentrations of 300 m*M*, 320 m*M*, and about 1 *M*, respectively (Hensel and König, 1988; Martins et al., 1997). cDPG at lower concentrations was detected in *Pyrococcus woesei*, and the available genome sequence data suggest

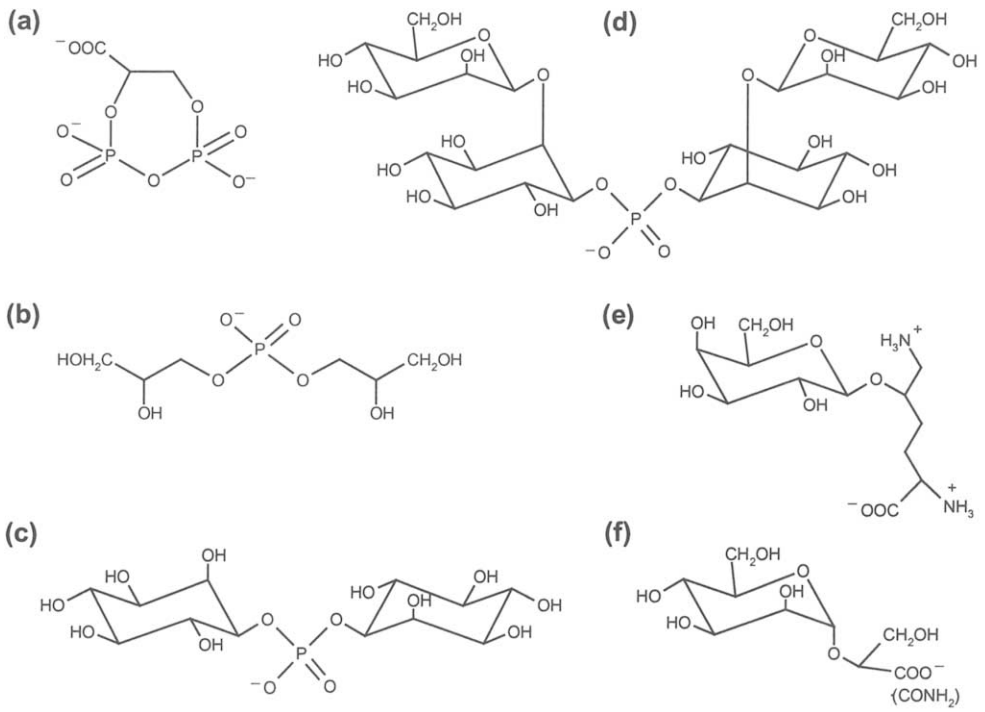


FIGURE 14. Compatible solutes from (hyper)thermophiles. (a) Cyclic 2,3-diphosphoglycerate, (b) diglycerol phosphate, (c) di-*myo*-inositol-1,1′-phosphate, (d) di-2-*O*-β-mannosyl-di-*myo*-inositol-1,1′-phosphate, (e) β-galactopyranosyl-5-hydroxylysine, (f) α-mannosylglycerate and α-mannosylglyceramide.

that *Pyrococcus horikoshii* is also able to synthesize this compound. Thus, cDPG production is not restricted to the methanogenic *Euryarchaeota* (Matussek et al., 1998). In *Methanothermus fervidus* the intracellular concentrations of the trianionic cDPG and of its potassium counter-cation increase with increasing temperature, up to the optimum growth temperature of 84°C (Hensel and König, 1988). *In vitro* experiments demonstrated that 300 m*M* potassium-cDPG increased the half-lives at 90°C of malate dehydrogenase and glyceraldehyde-3-phosphate dehydrogenase from *Methanothermus fervidus* from 8 min to 16 h and from 1.5 min to 3.5 h, respectively. In contrast, mesophilic rabbit glyceraldehyde-3-phosphate dehydrogenase was not stabilized by potassium-cDPG (Hensel and König, 1988). The rapid inactivation of formyltransferase from *Methanopyrus kandleri* in a low-salt buffer at 90°C was prevented by the addition of ≥0.5 *M* potassium-cDPG. Potassium phosphate at >0.7 *M* had the same stabilizing effect on the formyltransferase (Shima et al., 1998). Because most of the enzymes of methanogenesis of *Methanopyrus kandleri* are dependent on high concentrations of potassium phosphate or ammonium sulfate for activity and/or stability (Shima et al., 1998), it seems likely that these enzymes are also stabilized by potassium-cDPG *in vivo*. One reason why cDPG and not phosphate is present in molar concentrations in the cells may be that such a high phosphate concentration would affect the ATP-forming and ATP-utilizing reactions in the metabolism of the hyperthermophile. By increasing the phosphate concentration from the millimolar range to 1 *M*, the free energy change of ATP hydrolysis would decrease from about –50 kJ/mol to about –30 kJ/mol (Shima et al., 1998)

2. Diglycerol Phosphate

Diglycerol phosphate (1,1′-Diglyceryl phosphate, DGP; Figure 14b) is the main compatible solute found in *Archaeoglobus fulgidus* under conditions of salt stress. Its highest observed intracellular concentration was about 350 m*M*, which was measured in the presence of 4.5 *M* NaCl in the growth medium (Martins et al., 1997; Lamosa et al., 2000). DGP also appears to be a thermoprotectant, because its concentration increased about twofold when the growth temperature was shifted from 76°C to 87°C. However, the DGP concentration reached under heat shock stayed below the concentration reached under salt stress (Martins et al., 1997). DGP efficiently protected *Thermococcus litoralis* glutamate dehydrogenase from thermal inactivation. After an incubation at 95°C for 80 min, the residual activity dropped by 80% in the absence of stabilizing additives, but by only 15% in the presence of 100 m*M* DGP. Potassium phosphate at the same concentration also provided considerable, yet less efficient protection. A protective effect of DGP against heat inactivation at 50 to 60°C was also observed with rabbit muscle lactate dehydrogenase and yeast alcohol dehydrogenase (Lamosa et al., 2000). However, in these cases potassium phosphate had a similarly strong stabilizing effect as potassium-DGP.

3. Derivatives of myo-Inositol Phosphate

Various organic solutes derived from *myo*-inositol phosphate have been found in extreme thermophiles. Di-*myo*-inositol-1,1′(3,3′)-phosphate (Figure 14c) was detected in several hyperthermophilic archaea, including *Methanococcus igneus, Pyrococcus furiosus, Pyrococcus woesei, Pyrodictium occultum, Thermococcus celer,* and *Thermococcus stetteri* (Lamosa et al., 1998; Martins and Santos, 1995; Martins et al., 1997; Scholz et al., 1992), and in the hyperthermophilic bacterium *Aquifex pyrophilus* (Martins et al., 1996). In *Archaeoglobus fulgidus*, di-*myo*-inositol phosphate accumulated (in addition to DGP; see above) in response to a temperature up-shift (Martins et al., 1997). Also for *Methanococcus igneus*, *Pyrococcus furiosus*, *Thermococcus celer,* and *Thermococcus stetteri,* a clear correlation between temperature increase and solute accumulation was shown (Ciulla et al., 1994; Lamosa et al., 1998; Martins et al., 1995).

The hyperthermophilic bacteria *Thermotoga maritima* and *Thermotoga neapolitana* also contain di-*myo*-inositol-1,1′(3,3′)-phosphate, but additionally di-*myo*-inositol-1,3′-phosphate and di-2-*O*-β-mannosyl-di-*myo*-inositol-1,1′(3,3′)-phosphate (Figure 14d) (Martins et al., 1996). At optimum growth conditions with respect to temperature and salinity, these compounds were present at intracellular concentrations between 20 and 120 mM. They could not be detected in less thermophilic *Thermotogales* species. For *T. neapolitana* it was shown that salt stress (increase from 1.5 to 4.7%) as well as thermal stress (increase from 65°C to 88°C) resulted in significantly elevated levels of the three *myo*-inositol-phosphate derivatives mentioned before. Intriguingly, the concentrations of di-*myo*-inositol-1,3′-phosphate and di-2-*O*-β-mannosyl-di-*myo*-inositol-1,1′(3,3′)-phosphate, which are marginal at "low" temperatures of 65°C and 75°C, dramatically increase above the optimum growth temperature of 80°C and are the major compatible solutes in *T. neapolitana* between about 85°C and 90°C (Martins et al., 1996). Similarly, Ramakrishnan et al. (1997) reported the temperature-dependent increase of the intracellular di-*myo*-inositol-phosphate level in *Thermotoga maritima*. Little infor-

mation is available about specific protein thermoprotection *in vitro* by *myo*-inositol phosphate derivatives. Di-*myo*-inositol-phosphate was found to stabilize *Pyrococcus woesei* glyceraldehyde-3-phosphate dehydrogenase, but sodium citrate had a similar effect (Scholz et al., 1992). Di-*myo*-inositol-phosphate from *Thermotoga maritima* was not able to stabilize pyruvate ferredoxin oxidoreductase or hydrogenase of this organism against heat inactivation at 90°C (Ramakrishnan et al., 1997). Nevertheless, the occasionally observed heat-induced accumulation of *myo*-inositol phosphate derivatives suggests that these substances could be of importance *in vivo* to protect proteins from hyperthermophilic archaea and bacteria.

4. β-Galactopyranosyl-5-Hydroxylysine

β-galactopyranosyl-5-hydroxylysine (Figure 14e) is a compatible solute, which to date was detected only in the hyperthermophilic archaeon *Thermococcus litoralis* (Lamosa et al., 1998). The compound accumulated in response to increased salinity. Because β-galactopyranosyl-5-hydroxylysine was found only in cells grown in peptone-containing media, it probably derives from glycosyl-5-hydroxylysine moieties from collagen that are present in peptone preparations (Lamosa et al., 1998). Therefore, it is unlikely that β-galactopyranosyl-5-hydroxylysine is a specific thermoprotectant.

5. Mannosylglycerate and Mannosylglyceramide

α-Mannosylglycerate (Figure 14f) is a compatible solute found in the hyperthermophilic archaea *Pyrococcus furiosus, Pyrococcus woesei, Methanothermus fervidus* and *Thermococcus litoralis,* and in the moderately or extremely thermophilic bacteria *Petrotoga myotherma, Thermus thermophilus,* and *Rhodothermus marinus* (Lamosa et al., 1998; Martins et al., 1997; Nunes et al., 1995; Silva et al., 1999). At high salinity and low temperature, *Rhodothermus marinus* accumulated α-mannosylglyceramide (Figure 14f), but it was replaced at high temperature by α-mannosylglycerate (Silva et al., 1999). α-Mannosylglycerate, which has not been found in mesophilic bacteria, is formed mainly under salt stress (DaCosta et al., 1998), but its accumulation at elevated temperatures has also been reported (Silva et al., 1999). At 500 m*M* concentration, α-mannosylglycerate protected various thermostable and thermolabile enzymes *in vitro*, for example, alcohol dehydrogenase (but not glutamate dehydrogenase) from *Pyrococcus furiosus*, glutamate dehydrogenase from *Thermotoga maritima,* and rabbit muscle lactate dehydrogenase (Ramos et al., 1997). Because the α-mannosylglycerate concentration used for the *in vitro* experiments was, however, much higher than the intracellular concentrations typically reported for various organisms (Lamosa et al., 1998; Martins and Santos, 1995; Nunes et al., 1995; Silva et al., 1999), its protein protecting role *in vivo* remains questionable.

6. Protein and Species Specificity of Compatible Solutes

It has been observed that similar proteins are protected to a different extent by a given compatible solute. Along these lines, Lamosa et al. (2000) have carried

out a comparative study with three rubredoxins, which are structurally similar but are stabilzed differently by potassium-DGP, potassium phosphate, and glycerol. At 90°C, all tested compounds had only marginal effects on the stability of *Desulfovibrio desulfuricans* rubredoxin. *Clostridium pasteurianum* rubredoxin showed a fourfold increased half-life in the presence of 100 m*M* DGP, but was hardly stabilized by potassium phosphate. *Desulfovibrio gigas* rubredoxin showed a 3.5-fold increased half-life in the presence of 100 m*M* DGP, but was also stabilized more than twofold by potassium phosphate. Glycerol had no or only a very small protective effect in all three cases (Lamosa et al., 2000). Protein-specific differences in the thermoprotecting effect have also been found with other compatible solutes, for example, ectoine (Lippert and Galinski, 1992). Obviously, some proteins are sufficiently intrinsically stable, whereas others are more or less dependent on extrinsic stabilization to acquire the physiologically desirable half-lives *in vivo*.

In contrast to the compatible solute-producing thermophiles described in the preceding paragraphs, at least some thermophilic and hyperthermophilic organisms with a low salt requirement do not produce significant amounts of organic solutes at their optimum growth temperature. It was shown in a few cases that they fail to do so even after a temperature up-shift. For example, *Thermotoga thermarum, Fervidobacterium islandicum, Pyrobaculum islandicum,* and *Thermococcus zilligii* AN1 do not accumulate organic solutes during growth at 70°C, 70°C, 95°C, and 75 to 84°C, respectively (Martins et al., 1996, 1997; Lamosa et al., 1998). These organisms obviously do not need solutes as extrinsic protein stabilizers in the cell. Therefore, the precise *in vivo* roles of those compatible solutes of thermophiles and hyperthermophiles that are accumulated as the temperature is raised await further clarification. It has been speculated that some solutes such as DIP, cDPG, and mannosylglycerate, which are produced by slightly halophilic thermophiles, may function as thermoprotectants primarily when osmotic and high-temperature stress act simultaneously (Lamosa et al., 1998). However, this hypothesis has not been proven yet.

VIII. THERMOSTABLE ENZYMES IN BIOTECHNOLOGY

Enzymes are used in a large variety of applications in different fields of industry. Clearly, in some applications extremely thermostable enzymes would be very beneficial, for example, in applications where high process temperatures prevail (Niehaus et al., 1999). One prominent example is in the starch industry, where starch is gelatinized at elevated temperature to make the polysaccharide accessible for enzymatic attack. Thermostable enzymes of interest for starch conversion include amylases, pullulanase, glucoamylase, glucose isomerase, branching enzyme, and cyclodextrin glycosyltransferase (Niehaus et al., 1999; Zeikus et al., 1998). Another example is in the paper industry, which is increasingly aware of environmental risks caused by pollution of wastewater with hazardous chemicals. Therefore, paper manufacturers are aiming to reduce the amount of chlorine-containing chemicals needed for the bleaching of pulp. One of the strategies at trial is to remove residual xylan from the pulp with xylanases, which can significantly reduce the amount of the subsequently applied bleaching agents (Viikari et al., 1994). Xylanases useful for this

prebleaching process should optimally be active at elevated temperature (>70°C) and at alkaline pH.

Also in many other industrial applications extremely (thermo-)stable enzymes have general advantages over less stable catalysts. Running processes at high temperature will increase reaction rates, as it will (1) improve the solubility of substrates and products, (2) accelerate diffusion, (3) decrease viscosity, and (4) decrease the risk of microbial growth with its subsequent negative effects on the production machinery as well as product yield and quality. A highly advantageous aspect of extremely thermostable enzymes is their easy purification after recombinant production in mesophilic hosts, because simple heat treatment of a crude cell extract typically removes more than 90% of the heat-labile host proteins. Despite all the advantages the enzymes from hyperthermophiles have to offer, however, only one enzyme class characterized from these organisms has made it through to large-scale application in industry. This notable exception are thermostable DNA polymerases, which are used world-wide in academic and nonacademic laboratories for the polymerase chain reaction (PCR), one of the key methods of modern genetic engineering, DNA sequencing, and molecular diagnostics (Niehaus et al., 1999). The reasons for the failure of most of the hyperthermophiles' enzymes to make it to the market so far are multifacetted. Apart from performing well in terms of the desired catalytic activity in laboratory tests, it must be possible to produce the enzymes on a large scale at a low price. In addition, the enzymes need to fulfill a multitude of other criteria in order to successfully compete with established less thermostable enzymes that have been found in extensive screening programs during recent decades. For example, the applicability of an enzyme in a detergent product will depend on its performance at different temperatures, its stability and activity in the presence of a complex mixture of detergents and additives, and its shelf life in granular or liquid formulations, to name but a few criteria.

Nevertheless, the enzymes of extremely thermophilic organisms due to their robust nature remain highly interesting from an industrial point of view, but more research in this area is necessary. As only relatively few extreme thermophiles have been analyzed in depth, it is important to extend our knowledge about these organisms and their enzymes. The natural microbial diversity present in hot habitats needs to be examined and exploited more thoroughly in order to find novel and highly stable biocatalysts with the properties needed in biotechnological applications. In some commercial applications it would be useful to combine the high stability inherent in enzymes from extreme thermophiles with high catalytic activity at moderate temperatures. Along these lines, high thermostability and high catalytic activity at low temperature could be combined in industrially interesting proteases by directed evolution (Giver et al., 1998; Miyazaki and Arnold, 1999; Miyazaki et al., 2000; see Sections V and VI). The technology of *in vitro* evolution is expected to have a large impact on the development of industrial biocatalysts (Arnold and Volkov, 1999).

In conclusion, the various experimental approaches and results summarized in this review, including research in the fields of microbiology, molecular biology, biochemistry, biophysical chemistry, and structural biology, have revealed a number of mechanisms of thermophilic adaptation of proteins. Nevertheless, continued interdisciplinary efforts are necessary to develop a deep understanding of how hyperthermophilic proteins are stabilized

intrinsically and extrinsically, and to exploit this knowledge in biotechnology.

ACKNOWLEDGMENT

Work in the authors' laboratories is supported by the Deutsche Forschungsgemeinschaft, the Fonds der Chmischen Industrie, and the Deutsche Bundesstiftung Umwelt. We thank Profs. Kasper Kirschner and Franz X. Schmid for valuable comments on the manuscript, and Stefan Hettwer for preparing figures.

REFERENCES

Akanuma, S., Yamagishi, A., Tanaka, N., and Oshima, T. 1998. Serial increase in the thermal stability of 3-isopropylmalate dehydrogenase from *Bacillus subtilis* by experimental evolution. *Prot. Sci*. **7:** 698–705.

Akanuma, S., Yamagishi, A., Tanaka, N., and Oshima, T. 1999. Further improvement of the thermal stability of a partially stabilized *Bacillus subtilis* 3-isopropylmalate dehydrogenase variant by random and site-directed mutagenesis. *Eur. J. Biochem.* **260:** 499–504.

Andrä, S., Frey, G., Jaenicke, R., and Stetter, K. O. 1998. The thermosome from *Methanopyrus kandleri* possesses an NH_4^+-dependent ATPase activity. *Eur. J. Biochem.* **255:** 93–99.

Archibald, J. M., Logsdon, J. M. Jr., and Doolittle, W. F. 1999. Recurrent paralogy in the evolution of archaeal chaperonins. *Curr. Biol.* **9:** 1053–1056.

Argüelles, J. C. 2000. Physiological roles of trehalose in bacteria and yeasts: a comparative analysis. *Arch. Microbiol.* **174:** 217–224.

Arnold, F. H. and Volkov, A. A. 1999. Directed evolution of biocatalysts. *Curr. Opin. Chem. Biol.* **3**: 54–59.

Arnott, M. A., Michael, R. A., Thompson, C. R., Hough, D., and Danson, M. J. 2000. Thermostability and Thermoactivity of Citrate Synthases from the Thermophilic and Hyperthermophilic Archaea, *Thermoplasma acidophilum* and *Pyrococcus furiosus*. *J. Mol. Biol.* **304:** 657–668.

Auerbach, G., Ostendorp, R., Prade, L., Korndörfer, I., Dams, T., Huber, R., and Jaenicke, R. 1998. Lactate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: the crystal structure at 2.1 Å resolution reveals strategies for intrinsic protein stabilization. *Structure* **6:** 769–781.

Barns, S. M., Fundyga, R. E., Jeffries, M. W., and Pace, N. R. 1994. Remarkable archaeal diversity in a Yellowstone National Park hot spring environment. *Proc. Natl. Acad. Sci. USA* **91:** 1609–1613.

Barns, S. M., Delwiche, C. F., Palmer, J. D., and Pace, N. R. 1996. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc. Natl. Acad. Sci. USA* **93:** 9188–9193.

Blöchl, E., Rachel, R., Burggraf, S., Hafenbradl, D., Jannasch, H. W., and Stetter, K. O. 1997. *Pyrolobus fumarii*, gen. And sp. Nov., represents a novel group of archaea, extending the upper temperature limit for life to 113°C. *Extremophiles* **1:** 14–21.

Brennan, T. V., Anderson, J. W., Jia, Z., Waygood, E. B., and Clarke, S. 1994. Repair of spontaneously deamidated HPr phosphocarrier protein catalyzed by the L-isoaspartate-(D-aspartate) *O*-methyltransferase. *J. Biol. Chem.* **269:** 24586–24595.

Bukau, B. and Horwich, A. L. 1998. The Hsp70 and Hsp60 chaperone machines. *Cell* **92:** 351–366.

Burggraf, S., Heyder, P., and Eis, N. 1997. A pivotal Archaea group. *Nature* **385:** 780.

Cambillau, C. and Claverie, J.-M. 2000. Structural and genomic correlates of hyperthermostability. *J. Biol. Chem.* **275:** 32383–32386.

Cavagnero, S., Zhou, Z. H., Adams, M. W., and Chan, S. I. 1995. Response of rubredoxin from *Pyrococcus furiosus* to environmental changes: implications for the origin of hyperthermostabiliy. *Biochemistry* **34:** 9865–9873.

Cavagnero, S., Debe, D. A., Zhou, Z. H., Adams, M. W., and Chan, S. I. 1998. Kinetic role of electrostatic interactions in the unfolding of hyperthermophilic and mesophilic rubredoxins. *Biochemistry* **37:** 3369–3376.

Chakravarty, S. and Varadarajan, R. 2000. Elucidation of determinants of protein stability through genome sequence analysis. *FEBS Lett.* **470:** 65–69.

Ciulla, R. A., Burggraf, S., Stetter, K. O., and Roberts, M. F. 1994. Occurrence and role of di-*myo*-inositol-1,1'-phosphate in *Methanococcus igneus*. *Appl. Environ. Microbiol.* **60:** 3660–3664.

Crameri, A., Raillard, S. A., Bermudez, E., and Stemmer, W. P. 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391:** 288–291.

DaCosta, M. S., Santos, H., and Galinski, E. A. 1998. An overview of the role and diversity of compatible solutes in *Bacteria* and *Archaea*. In: *Biotechnology of Extremophiles*, pp. 118–153 (Antranikian, G., Ed.), Springer-Verlag, Berlin.

Dams, T. and Jaenicke, R. 1999. Stability and folding of dihydrofolate reductase from the hyperthermophilic bacterium *Thermotoga maritima*. *Biochemistry* **38:** 9169–9178.

Daniel, R. M., Dines, M., and Petach, H. H. 1996. The denaturation and degradation of stable enzymes at high temperatures. *Biochem. J.* **317:** 1–11.

Dao-Pin, S., Sauer, U., Nicholson, H., and Matthews, B. W. 1991. Contributions of engineered surface salt bridges to the stability of T4 lysozyme determined by directed mutagenesis. *Biochemistry* **30:** 7142–7153.

Darimont, B. 1994. Studies on Catalysis, Folding and Evolution of Indoleglycerol Phosphate Synthase, an Eightfold βα Barrel Involved in Tryptophan Biosynthesis. Ph.D. thesis, University of Basel, Switzerland.

Darimont, B., Stehlin, C., Szadkowski, H., and Kirschner, K. 1998. Mutational analysis of the active site of indoleglycerol phosphate synthase from *Escherichia coli*. *Prot. Sci.* **7:** 1221–1232.

De Bakker, P. I., Hunenberger, P. H., and McCammon, J. A. 1999. Molecular dynamics simulations of the hyperthermophilic protein sac7d from *Sulfolobus acidocaldarius*: contribution of salt bridges to thermostability. *J. Mol. Biol.* **285:** 1811–1830.

Deckert, G., Warren, P.V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J., and Swanson, R. V. 1998. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392:** 353–358.

Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., and Bukau, B. 1999. Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature* **400:** 693–696.

Ditzel, L., Lowe, J., Stock, D., Stetter, K. O., Huber, H., Huber, R., and Steinbacher, S. 1998. Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. *Cell* **93:** 125–138.

Elcock, A. H. 1998. The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *J. Mol. Biol.* **284:** 489–502.

Elcock, A. H. and McCammon, J. A. 1997. Continuum solvation model for studying protein hydration thermodynamics at high temperatures. *J. Phys. Chem.* **101:** 9624–9634.

Emmerhoff, O. J., Klenk, H. P., and Birkeland, N. K. 1998. Characterization and sequence

comparison of temperature-regulated chaperonins from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *Gene* **215:** 431–438.

Fabry, S. and Hensel, R. 1987. Purification and characterization of D-glyceraldehyde-3–phosphate dehydrogenase from the thermophilic archaebacterium *Methanothermus fervidus*. *Eur. J. Biochem.* **165:** 147–155.

Facchiano, A. M., Colonna, G., and Ragone, R. 1998. Helix stabilizing factors and stabilization of thermophilic proteins: an X-ray based study. *Prot. Eng.* **11:** 753–760.

Furutani, M., Iida, T., Yoshida, T., and Maruyama, T. 1998. Group II chaperonin in a thermophilic methanogen, *Methanococcus thermolithotrophicus* — chaperone activity and filament-forming ability. *J. Biol. Chem.* **273:** 28399–28407.

Geiger, T. and Clarke, S. 1987. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *J. Biol. Chem.* **262:** 785–794.

Giver, L., Gershenson, A., Freskgard, P. O., and Arnold, F. H. 1998. Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. U.S.A.* **95:** 12809–12813.

Glover, J. R. and Lindquist, S. 1998. Hsp104, Hsp70 and Hsp40: a novel chaperone system that rescues previously aggregated proteins. *Cell* **94:** 73–82.

Goloubinoff, P., Mogk, A., Peres Ben Zvi, A., Tomoyasu, T., and Bukau, B. 1999. Sequential mechanism of solubilisation and refolding of stable protein aggregates by a bichaperone network. *Proc. Natl. Acad. Sci. USA* **96:** 13732–13737.

Gribaldo, S., Lumia, V., Creti, R., de Macario, E. C., Sanangelantoni, A., and Cammarano, P. 1999. Discontinuous occurrence of the hsp70 (dnaK) gene among *Archaea* and sequence features of HSP70 suggest a novel outlook on phylogenies inferrred from this protein. *J. Bacteriol.* **181:** 434–443.

Grimsley, G. R., Shaw, K. L., Fee, L. R., Alston, R. W., Huyghues-Despointes, B. M., Thurlkill, R. L., Scholtz, J. M., and Pace, C. N. 1999. Increasing protein stability by altering long-range coulombic interactions. *Prot. Sci.* **8:** 1843–1849.

Grogan, D. W. 1998. Hyperthermophiles and the problem of DNA instability. *Mol. Microbiol.* **28:** 1043–1049.

Gutsche, I., Essen, L.-O., and Baumeister, W. 1999. Group II chaperonins: new TriC(k)s and turns of a protein folding machine. *J. Mol. Biol.* **293:** 295–312.

Gutsche, I., Mihalache, O., and Baumeister, W. 2000a. ATPase cycle of an archaeal chaperonin. *J. Mol. Biol.* **300:** 187–196.

Gutsche, I., Holzinger, J., Roßle, M., Heumann, H., Baumeister, W., and May, R. P. 2000b. Conformational rearrangements of an archaeal chaperonin upon ATPase cycling. *Curr. Biol.* **10:** 405–408.

Gutsche, I., Mihalache, O., Hegerl, R., Typke, D., and Baumeister, W. 2000c. ATPase cycle controls the conformation of an archaeal chaperonin as visualized by cryo-electron microscopy. *FEBS Lett.* **477:** 278–282.

Haley, D. A., Bova, M. P., Huang, Q.-L., Michaourab, H. S., and Stewart, P. L. 2000. Small heat-shock protein structures reveal a continuum from symmetric to variable assemblies. *J. Mol. Biol.* **298:** 261–272.

Haney, P. J., Badger, J. H., Buldak, G. L., Reich, C. I., Woese, C. R., and Olsen, G. J. 1999. Thermal adaptation analyzed by comparison of protein sequences from mesophilic and extremely thermophilic *Methanococcus* species. *Proc. Natl. Acad. Sci. USA* **96:** 3578–3583.

Hayes, C. S. and Setlow, P. 1997. Analysis of deamidation of small, acid-soluble spore proteins from *Bacillus subtilis in vitro* and *in vivo*. *J. Bacteriol.* **179:** 6020–6027.

Hennig, M., Darimont, B., Sterner, R., Kirschner, K., and Jansonius, J. N. 1995. 2.0 Å structure of indole-3–glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability *Structure* **3:** 1295–1306.

Hennig, M., Sterner, R., Kirschner, K., and Jansonius, J. N. 1997. Crystal structure at 2.0 Å resolution of phosphoribosylanthranilate isomerase from the hyperthermophile *Thermotoga maritima*: possible determinants of protein stability. *Biochemistry* **36:** 6009–6016.

Hensel, R. and Jakob, I. 1994. Stability of glyceraldehyde-3–phosphate dehydrogenases from hyperthermophilic archaea at high temperature. System. *Appl. Microbiol.* **16:** 742–745.

Hensel, R. and König, H. 1988. Thermoadaptation of methanogenic bacteria by intracellular ion concentration. *FEMS Microbiol. Lett.* **49:** 75–79.

Hernandez, G., Jenney, F. E., Jr., Adams, M. W., and LeMaster, D. M. 2000. Millisecond time scale conformational flexibility in a hyperthermophile protein at ambient temperature. *Proc. Natl. Acad. Sci. USA* **97:** 3166–3170.

Hesterkamp, T. and Bukau, B. 1998. Role of the DnaK and HscA homologs of Hsp70 chaperones in protein folding in *E. coli*. *EMBO J.* **17:** 4818–4828.

Hiller, R., Zhou, Z. H., Adams, M. W., and Englander, S. W. 1997. Stability and dynamics in a hyperthermophilic protein with melting temperature close to 200°C. *Proc. Natl. Acad. Sci. USA* **94:** 11329–11332.

Hollien, J. and Marqusee, S. 1999. A thermodynamic comparison of mesophilic and thermophilic ribonucleases H. *Biochemistry* **38:** 3831–3836.

Horwich, A. L., Weber-Ban, E. U., and Finley, D. 1999. Chaperone rings in protein folding and degradation. *Proc. Natl. Acad. Sci. USA* **96:** 11033–11040.

Hoseki, J., Yano, T., Koyama, Y., Kuramitsu, S., and Kagamiyama, H. 1999. Directed evolution of thermostable kanamycin-resistance gene: a convenient selection marker for *Thermus thermophilus*. *J. Biochem.* **126:** 951–956.

Hough, D. W. and Danson, M. J. 1999. Extremozymes. *Curr. Opin. Chem. Biol.* **3:** 39–46.

Huber, R., Burggraf, S., Mayer, T., Barns, S. M., Roßnagel, P., and Stetter, K. O. 1995. Isolation of a hyperthermophilic archaeum predicted by *in situ* RNA analysis. *Nature* **376:** 57–58.

Hugenholtz, P., Pitulle, C., Herschberger, K. L., and Pace, N. R. 1998. Novel division level bacterial diversity in a Yellowstone hot spring. *J. Bacteriol.* **180:** 366–376.

Ibarra-Molero, B., Loladze, V,V,, Makhatadze, G. I., and Sanchez-Ruiz, J. M. 1999. Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. *Biochemistry* **38:** 8138–8149.

Ichikawa, J. K. and Clarke, S. 1998. A highly active protein repair enzyme from an extreme thermophile: the L-isoaspartyl methyltransferase from *Thermotoga maritima*. *Arch. Biochem. Biophys.* **358:** 222–231.

Jaenicke, R. 1991. Protein stability and molecular adaptations to extreme conditions. *Eur. J. Biochem.* **202:** 715–728.

Jaenicke, R. 1998. What ultrastable globular proteins teach us about protein stability. *Biochemistry (Moscow)* **63:** 312–321.

Jaenicke, R. 2000. Do ultrastable proteins from hyperthermophiles have high or low conformational rigidity? *Proc. Natl. Acad. Sci. USA* **97:** 2962–2964.

Jaenicke, R. and Böhm, G. 1998. The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* **8:** 738–748.

Jaenicke, R., Schurig, H., Beaucamp, N., and Ostendorp, R. 1996. Structure and stabil-

ity of hyperstable proteins: glycolytic enzymes from the hyperthermophilic bacterium *Thermotoga maritima*. *Adv. Prot. Chem.* **48:** 181–269.

Joachimiak, A., Quaite-Randall, E., Tollaksen, S., Mai, X., Adams, M. W. M., Josephs, R., and Giomett, C. 1997. Purification of chaperonins from thermophilic bacteria and archaea. *J. Chromat. Ser. A* **773:** 131–138.

Johnson, B. A., Murray, E. D., Jr., Clarke, S., Glass, D. B., and Aswad, D. W. 1987. Protein carboxyl methyltransferase facilitates conversion of atypical L-isoaspartyl peptides to normal L-aspartyl peptides. *J. Biol. Chem.* **262:** 5622–5629.

Kagawa, H. K. Osipiuk, J., Maltsev, N., Overbeek, R., Quaite-Randall, E., Joachimiak, A., and Trent, J. D. 1995. The 60 kDa heat shock proteins in the hyperthermophilic archaeon *Sulfolobus shibatae*. *J. Mol. Biol.* **253:** 712–725.

Kanodia, S., and Roberts, M. F. 1983. Methanophosphagen: unique cyclic pyrophosphate isolated from *Methanobacterium thermoautotrophicum*. *Proc. Natl. Acad. Sci. USA* **80:** 5217–5221.

Karshikoff, A. and Ladenstein, R. 1998. Proteins from thermophilic and mesophilic organisms essentially do not differ in packing. *Protein Eng.* **11:** 867–872.

Kengen, S. W. M., de Bok, F. A. M., van Loo, N. D., Dijkema, C., Stams, A. J. M., and de Vos, W. M. 1994. Evidence for the operation of a novel Embden-Meyerhof pathway that involves ADP-dependent kinases during sugar fermentation by *Pyrococcus furiosus*. *J. Biol. Chem.* **269:** 17537–17541.

Kengen, S. W. M., Stams, A. J. M., and de Vos, W. M. 1996. Sugar metabolism of hyperthermophiles. *FEMS Microbiol. Rev.* **18:** 119–137.

Kim, K. K., Kim, R., and Kim, S.-H. 1998a. Crystal structure of a small heat-shock protein. *Nature* **394:** 595–599.

Kim, R., Kim, K. K., and Kim, S.-H. 1998b. Small heat-shock protein of *Methanococcus jannaschii*, a hyperthermophile. *Proc. Natl. Acad. Sci. USA* **95:** 9129–9133.

Klostermeier, D., Seidel, R., and Reinstein, J. 1999. The functional cycle and regulation of the *Thermus thermophilus* DnaK chaperone system. *J. Mol. Biol.* **287:** 511–525.

Klumpp, M., Baumeister, W., and Essen, L. O. 1997. Structure of the substrate binding domain of the thermosome, an archaeal group II chaperonin. *Cell* **91:** 263–270.

Knapp, S., Karshikoff, A., Berndt, K. D., Christova, P., Atanasov, B., and Ladenstein, R., 1996. Thermal unfolding of the DNA-binding protein Sso7d from the hyperthermophile *Sulfolobus solfataricus*. *J. Mol. Biol.* **264:** 1132–1144.

Knapp, S., Ladenstein, R., and Galinski, E. A. 1999. Extrinsic protein stabilization by the naturally occurring osmolytes β-hydroxyectoine and betaine. *Extremophiles* **3:** 191–198.

Koch, R., Spreinat, A., Lemke, K., and Antranikian, G. 1991. Purification and properties of a α-amylase from the archaeobacterium *Pyrococcus woesei*. *Arch. Microbiol.* **155:** 572–578.

Kohlhoff, M., Dahm, A., and Hensel, R. 1996. Tetrameric triosephosphate isomerase from hyperthermophilic Archaea. *FEBS Lett.* **383:** 245–250.

Korndörfer, I., Steipe, B., Huber, R., Tomschy, A., and Jaenicke, R. 1995. The crystal structure of holo-glyceraldehyde-3–phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima* at 2.5 Å resolution. *J. Mol. Biol*. **246:** 511–521.

Kowalski, J. M., Kelly, R. M., Konisky, J., Clark, D. S., and Wittrup, K. D. 1998. Purification and functional characterization of a chaperone from *Methanococcus jannaschii*. *Syst. Appl. Microbiol.* **21:** 173–178.

Kumar, S., Tsai, C. J., and Nussinov, R. 2000a. Factors enhancing protein thermostability. *Prot. Eng.* **13:** 179–191.

Kumar, S., Ma, B., Tsai, C. J., and Nussinov, R. 2000b. Electrostatic strengths of salt bridges in thermophilic and mesophilic glutamate dehydrogenase monomers. *Proteins* **38:** 368–383.

Lamosa, P., Martins, L. O., DaCosta, M. S., and Santos, H. 1998. Effects of temperature, salinity, and medium composition on compatible solute accumulation by *Thermococcus* spp. *Appl. Environ. Microbiol.* **64:** 3591–3598.

Lamosa, P., Burke, A., Peist, R., Huber, R., Liu, M.-Y., Silva, G., Rodrigues-Pousada, C., LeGall, J., Maycock, C., and Santos, H. 2000. Thermostabilization of proteins by diglycerol phosphate, a new compatible solute from the hyperthermophile *Archaeoglobus fulgidus*. *Appl. Environ. Microbiol.* **66:** 1974–1979.

Lebbink, J. H., Kaper, T., Bron, P., van der Oost, J., and de Vos, W. M. 2000. Improving low-temperature catalysis in the hyperthermostable *Pyrococcus furiosus* β-glucosidase CelB by directed evolution. *Biochemistry* **39:** 3656–3665.

Lehmann, M., Pasamontes, L., Lassen, S.F., and Wyss, M. 2000. The consensus concept for thermostability engineering of proteins. *Biochem. Biophys. Acta* **1543:** 408–415.

Leibrock, E., Bayer, P., and Lüdemann, H.-D. 1995. Non-enzymatic hydrolysis of ATP at high temperatures and high pressures. *Biophys. Chem.* **54:** 175–180.

Leroux, M. P. and Hartl, F. U. 2000. Protein folding: Versatility of the cytosolic chaperonin TriC/CCT. *Curr. Biol.* **10:** R260–R264.

Leroux, M. P., Fändrich, M., Klunker, D., Siegers, K., Lupas, A. N., Brown, J. R., Schiebel, E., Dobson, C. M., and Hartl, F. U. 1999. MtGimC, a novel archaeal chaperone related to the eukaryotic chaperonin cofactor GimC/prefoldin. *EMBO J.* **18:** 6730–6743.

Li, C. and Clarke, S. 1992. A protein methyltransferase specific for altered aspartyl residues is important in *Escherichia coli* stationary-phase survival and heat-shock resistance. *Proc. Natl. Acad. Sci. USA* **89:** 9885–9889.

Li, W. T., Grayling, R. A., Sandman, K., Edmondson, S., Shriver, J. W., and Reeve, J. N. 1998. Thermodynamic stability of archaeal histones. *Biochemistry* **37:** 10563–10572.

Li, W. T., Shriver, J. W., and Reeve, J. N. 2000. Mutational analysis of differences in thermostability between histones from mesophilic and hyperthermophilic archaea. *J. Bacteriol.* **182:** 812–817.

Lippert, K. and Galinski, E. A. 1992. Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. *Appl. Microbiol. Biotechnol.* **37:** 61–65.

Loladze, V. V., Ibarra-Molero, B., Sanchez-Ruiz, J. M., and Makhatadze, G. I. 1999. Engineering a thermostable protein via optimization of charge-charge interactions on the protein surface. *Biochemistry* **38:** 16419–16423.

Lorimer, G. H. 1997. Folding with a two-stroke motor. *Nature* **388:** 720–723.

Macario, A. J. L., Lange, M., Ahring, B. K., and Conway de Macario, E. 1999. Stress genes and proteins in the archaea. *Microbiol. Mol. Biol. Rev.* **63:** 923–967.

Maes, D., Zeelen, J. P., Thanki, N., Beaucamp, N., Alvarez, M., Thi, M. H., Backmann, J., Martial, J. A., Wyns, L., Jaenicke, R., and Wierenga, R. K. 1999. The crystal structure of triosephosphate isomerase (TIM) from *Thermotoga maritima*: a comparative thermostability structural analysis of ten different TIM structures. *Proteins* **37:** 441–453.

Makhatadze, G. I. and Privalov, P. L. 1995. Energetics of protein structure. *Adv. Protein Chem.* **47:** 307–425.

Malakauskas, S. M. and Mayo, S. L. 1998. Design, structure and stability of a hyperthermophilic protein variant. *Nat. Struct. Biol*. **5:** 470–475.

Malin, G. and Lapidot, A. 1996. Induction of synthesis of tetrahydropyrimidine derivatives in *Streptomyces* strains and their effect on *Escherichia coli* in response to osmotic and heat stress. *J. Bacteriol.* **178:** 385–395.

Martins, L. O. and Santos, H. 1995. Accumulation of mannosylglycerate and di-*myo*-inositol-phosphate by *Pyrococcus furiosus* in response to salinity and temperature. *Appl. Environ. Microbiol.* **61:** 3299–3303.

Martins, L. O., Carreto, L. S., Da Costa, M. S., and Santos, H. 1996. New compatible solutes related to di-*myo*-inositol-phosphate in members of the order *Thermotogales*. *J. Bacteriol.* **178:** 5644–5651.

Martins, L. O., Huber, R., Huber, H., Stetter, K. O., DaCosta, M. S., and Santos, H. 1997. Organic solutes in hyperthermophilic archaea. *Appl. Environ. Microbiol.* **63:** 896–902.

Matussek, K., Moritz, P., Brunner, N., Eckerskorn, C., and Hensel, R. 1998. Cloning, sequencing, and expression of the gene encoding cyclic 2,3–diphosphoglycerate synthetase, the key enzyme of cyclic 2,3–diphosphoglycerate metabolism in *Methanothermus fervidus*. *J. Bacteriol.* **180:** 5997–6004.

McCrary, B. S., Edmondson, S. P., and Shriver, J. W. 1996. Hyperthermophile protein folding thermodynamics: differential scanning calorimetry and chemical denaturation of Sac7d. *J. Mol. Biol.* **264:** 784–805.

Merz, A., Knöchel, T., Jansonius, J. N., and Kirschner, K. 1999. The hyperthermostable indoleglycerol phosphate synthase from *Thermotoga maritima* is destabilized by mutational disruption of two solvent-exposed salt bridges. *J. Mol. Biol.* **288:** 753–763.

Merz, A., Yee, M. C., Szadkowski, H., Pappenberger, G., Crameri, A., Stemmer, W. P., Yanofsky, C., and Kirschner, K. 2000. Improving the catalytic activity of a thermophilic enzyme at low temperatures. *Biochemistry* **39:** 880–889.

Minuth, T., Frey, G., Lindner, P., Rachel, R., Stetter, K. O., and Jaenicke, R. 1998. Recombinant homo- and hetero-oligomers of an ultrastable chaperonin from the archaeon *Pyrodictium occultum* show chaperone activity *in vitro*. *Eur. J. Biochem.* **258:** 837–845.

Minuth, T., Henn, M., Rutkat, K., Andrä, S., Frey, G., Rachel, R., Stetter, K. O., and Jaenicke, R. 1999. The recombinant thermosome from the archaeon *Methanopyrus kandleri*: *in vitro* analysis of its chaperone activity. *Biol. Chem.* **380:** 55–62.

Miyazaki, K. and Arnold, F. H. 1999. Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. *J. Mol. Evol.* **49:** 716–720.

Miyazaki, K., Wintrode, P. L., Grayling, R. A., Rubingh, D. N., and Arnold, F. H. 2000. Directed evolution study of temperature adaptation in a psychrophilic enzyme. *J. Mol. Biol.* **297:** 1015–1026.

Mogk, A., Tomoyasu, T., Goloubinoff, P., Rüdiger, S., Röder, D., Langen, H., and Bukau, B. 1999. Identification of thermolabile *Escherichia coli* proteins: prevention and reversion of aggregation by DnaK and ClpB. *EMBO J.* **18:** 6934–6949.

Motohashi, K., Watanabe, Y., Yohda, M., and Yoshida, M. 1999. Heat-inactivated proteins are rescued by the DnaK·J-GrpE set and ClpB chaperones. *Proc. Natl. Acad. Sci. USA* **96:** 7184–7189.

Mueller, U., Perl, D., Schmid, F. X., and Heinemann, U. 2000. Thermal stability and atomic-resolution crystal structure of the *Bacillus caldolyticus* cold shock protein. *J. Mol. Biol.* **297:** 975–988.

Murphy, K. P. and Freire, E. 1992. Thermodynamics of structural stability and coopera-

tive folding behavior in proteins. *Adv. Protein Chem.* **43:** 313–361.

Myers, J. K., Pace, C. N., and Scholtz, J. M. 1995. Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Prot. Sci.* **4:** 2138–2148.

Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum, K. A., McDonald, L., Utterback, T. R., Malek, J. A., Linher, K. D., Garrett, M. M., Stewart, A. M., Cotton, M. D., Pratt, M. S., Phillips, C. A., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann, R. D., Eisen, J. A., White, O., Salzberg, S. L., Smith, H. O., Venter, J. C., and Fraser, C. M. 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399:** 323–329.

Nemeth, A., Svingor, A., Pocsik, M., Dobo, J., Magyar, C., Szilagyi, A., Gal, P., and Závodszky, P. (2000). Mirror image mutations reveal the significance of an intersubunit ion cluster in the stability of 3–isopropylmalate dehydrogenase. *FEBS Lett.* **468:** 48–52.

Netzer, W. J. and Hartl, F. U. 1998. Protein folding in the cytosol: chaperonin-dependent and –independent mechanisms. *Trends Biochem. Sci.* **23:** 68–73.

Niehaus, F., Bertoldo, C., Kähler, M., and Antranikian, G. 1999. Extremophiles as a source of novel enzymes for industrial application. *Appl. Microbiol. Biotechnol.* **51:** 711–729.

Nunes, O. C., Manaia, C. M., Da Costa, M. S., and Santos, H. 1995. Compatible solutes in the thermophilic bacteria *Rhodothermus marinus* and *Thermus thermophilus*. *Appl. Environ. Microbiol.* **61:** 2351–2357.

Pace, C. N. 2000. Single surface stabilizer. *Nat. Struct. Biol*. **7:** 345–346.

Pace, C. N. and Scholtz, J. M. 1997. Measuring the conformational stability of a protein. In: *Protein structure-a practical approach*, pp. 299–321. (Creighton, T. E., Ed.). Oxford University Press, Oxford, New York.

Pappenberger, G., Schurig, H., and Jaenicke, R. 1997. Disruption of an ionic network leads to accelerated thermal denaturation of D-glyceraldehyde-3–phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*. *J. Mol. Biol*. **274:** 676–683.

Perl, D., Welker, C., Schindler, T., Schröder, K., Marahiel, M. A., Jaenicke, R., and Schmid, F. X. 1998. Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nat. Struct. Biol.* **5:** 229–235.

Perl, D., Mueller, U., Heinemann, U., and Schmid, F. X. 2000. Two exposed amino acid residues confer thermostability on a cold shock protein *Nat. Struct. Biol.* **7:** 380–383.

Perutz, M. F. and Raidt, H. 1975. Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2. *Nature* **255:** 256–259.

Petukhov, M., Kil, Y., Kuramitsu, S., and Lanzov, V. 1997. Insights into thermal resistance of proteins from the intrinsic stability of their alpha-helices. *Proteins* **29:** 309–320.

Pfeil, W. 1998. *Protein Stability and Folding. A Collection of Thermodynamic Data*, pp. 3–14. Springer-Verlag, Berlin.

Pfeil, W., Gesierich, U., Kleemann, G. R., and Sterner, R. 1997. Ferredoxin of the hyperthermophile *Thermotoga maritima* is stable beyond the boiling point of water. *J. Mol. Biol.* **272:** 591–596.

Phipps, B. M., Typke, D., Hegerl, R., Volker, S., Hoffmann, A., Stetter, K. O., and Baumeister, W. 1993. Structure of a chaperone from a thermophilic archaebacterium. *Nature* **361:** 475–477.

Plaza del Pino, I. M., Ibarra-Molero, B., and Sanchez-Ruiz, J. M. 2000. Lower kinetic limit to protein thermal stability: a proposal regarding protein stability *in vivo* and its relation with misfolding diseases. *Proteins* **40:** 58–70.

Privalov, P.L. 1979. Stability of proteins: small globular proteins. *Adv. Prot. Chem.* **33:** 167–241.

Privalov, P.L. and Khechinashvili, N.N. 1974. A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. *J. Mol. Biol.* **86:** 665–684.

Quaite-Randall, E., Trent, J. D., Josephs, R., and Joachimiak, A. 1995. Conformational cycle of the archaeosome, a TCP1–like chaperonin from *Sulfolobus shibatae. J. Biol. Chem.* **270:** 28818–28823.

Ramakrishnan, V., Verhagen, M. F. J. M., and Adams, M. W. W. 1997. Characterization of di-*myo*-inositol-1,1′-phosphate in the hyperthermophilic bacterium *Thermotoga maritima*. *Appl. Environ. Microbiol.* **63:** 347–350.

Ramos, A., Raven, N. D. H., Sharp, R. J., Bartolucci, S., Rossi, M., Cannio, R., Lebbink, J., van der Oost, J., de Vos, W. M., and Santos, H. 1997. Stabilization of enzymes against thermal stress and freeze-drying by mannosylglycerate. *Appl. Environ. Microbiol.* **63:** 4020–4025.

Reysenbach, A.-L., Ehringer, M., and Hershberger, K. 2000. Microbial diversity at 83°C in Calcite springs, Yellowstone National Park: another environment where the *Aquificiales* and "Korarchaeota" coexist. *Extremophiles* **4:** 61–67.

Roseman, A. M., Chen, S., White, H., Braig, K., and Saibil, H. R. 1996. The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL. *Cell* **87:** 241–251.

Rüdiger, S., Buchberger, A., and Bukau, B. 1997. Interaction of Hsp70 chaperones with substrates. *Nature Struct. Biol.* **4:** 342–349.

Schmid, F. X. 1997. Optical spectroscopy to characterize protein conformation and conformational changes. In: *Protein Structure—A Practical Approach*, pp. 261–297. (Creighton, T. E., Ed.). Oxford University Press, Oxford.

Schoen, G., Quiate-Randall, E., Jiminez, J. L., Joachimiak, A., and Saibil, H. R. 2000. Three conformations of an archaeal chaperonin, TF55 from *Sulfolobus shibatae*. *J. Mol. Biol.* **269:** 813–819.

Scholz, S., Sonnenbichler, J., Schäfer, W., and Hensel, R. 1992. Di-*myo*-insitol-1,1′-phosphate: a new inositol phosphate isolated from *Pyrococcus woesei*. *FEBS Lett.* **306:** 239–242.

Schröder, H., Langer, T., Hartl, F. U., and Bukau, B. 1993. DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. *EMBO J.* **12:** 4137–4144.

Seely, R. J. and Fahrney, D. E. 1983. A novel diphospho-P,P′-diester from *Methanobacterium thermoautotrophicum*. *J. Biol. Chem.* **258:** 10835–10838.

Shibuya, H., Kaneko, S., and Hayashi, K. 2000. Enhancement of the thermostability and hydrolytic activity of xylanase by random gene shuffling. *Biochem. J.* **349:** 651–656.

Shima, S., Herault, D. A., Berkessel, A., and Thauer, R. K. 1998. Activation and thermostabilization effects of cyclic 2,3–diphosphoglycerate on enzymes from the hyperthermophilic *Methanopyrus kandleri*. *Arch. Microbiol.* **170:** 469–472.

Shima, S., Thauer, R. K., Ermler, U., Durchschlag, H., Tziatzios, C., and Schubert, D. 2000. A mutation affecting the association equilibrium of formyltransferase from the hyperthermophilic *Methanopyrus kandleri* and its influence on the enzyme's activity and thermostability. *Eur. J. Biochem.* **267:** 6619–6623.

Shoichet, B. K., Baase, W. A., Kuroki, R., and Matthews, B. W. 1995. A relationship

between protein stability and protein function. *Proc. Natl. Acad. Sci. USA* **92:** 452–456.

Shtilerman, M., Lorimer, G. H., and Englander, S. W. 1999. Chaperonin function: folding by forced unfolding. *Science* **284:** 822–825.

Siegers, K., Waldmann, T., Leroux, M. R., Grein, K., Shevchenko, A., Schiebel, E., and Hartl, F. U. 1999. Compartmentation of protein folding *in vivo*: sequestration of non-native polypeptide by the chaperonin-GimC system. *EMBO J.* **18:** 75–84.

Siegert, R., Leroux, M. R., Scheuffler, C., Hartl, F. U., and Moarefi, I. 2000. Structure of the molecular chaperone prefoldin: unique interaction of multiple coiled coil tentacles with unfolded proteins. *Cell* **103:** 621–623.

Silva, Z., Borges, N., Martins, L. O., Wait, R., DaCosta, M. S., and Santos, H. 1999. Combined effect of the growth temperature and salinity of the medium on the accumulation of compatible solutes by *Rhodothermus marinus* and *Rhodothermus obamensis*. *Extremophiles* **3:** 163–172.

Slobodkin, A. I., Jeanthon, C., L'Haridon, S., Nazina, T., Miroshnichenko, M., and Bonch-Osmolovskaya, E. 1999. Dissimilatory reduction of Fe(III) by thermophilic bacteria and archaea in deep subsurface petroleum reservoirs of Western Siberia. *Curr. Microbiol.* **39:** 99–102.

Song, J. K. and Rhee, J. S. 2000. Simultaneous enhancement of thermostability and catalytic activity of phospholipase A(1) by evolutionary molecular engineering. *Appl. Environ. Microbiol.* **66:** 890–894.

Spector, S., Wang, M., Carp, S. A., Robblee, J., Hendsch, Z. S., Fairman, R., Tidor, B., and Raleigh, D. P. 2000. Rational modification of protein stability by the mutation of charged surface residues. *Biochemistry* **39:** 872–879.

Spiller, B., Gershenson, A., Arnold, F. H., and Stevens, R. C. 1999. A structural view of evolutionary divergence. *Proc. Natl. Acad. Sci. USA* **96:** 12305–12310.

Stemmer, W. P. 1994a. DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **91:** 10747–10751.

Stemmer, W. P. 1994b. Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* **370:** 389–391.

Stephenson, R. C. and Clarke, S. 1989. Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. *J. Biol. Chem.* **264:** 6164–6170.

Sterner, R., Kleemann, G. R., Szadkowski, H., Lustig, A., Hennig, M., and Kirschner, K. 1996. Phosphoribosyl anthranilate isomerase from *Thermotoga maritima* is an extremely stable and active homodimer. *Prot. Sci.* **5:** 2000–2008.

Stetter, K. O. 1996. Hyperthermophilic prokaryotes. *FEMS Microbiol. Rev.* **18:** 149–158.

Stetter, K. O. 1999. Extremophiles and their adaptation to hot environments. *FEBS Lett.* **452:** 22–25.

Stetter, K. O., Huber, R., Blöchl, E., Kurr, M., Eden, R. D., Fiedler, M., Cash, H., and Vance, I. 1993. Hyperthermophilic archaea are thriving in deep North Sea and Alaskan oil reservoirs. *Nature* **300:** 743–745.

Strop, P. and Mayo, S. L. 2000. Contribution of surface salt bridges to protein stability. *Biochemistry* **39:** 1251–1255.

Strop, P., Marinescu, A.M., and Mayo, S.L. 2000. Structure of a protein G helix variant suggests the importance of helix propensity and helix dipole interactions in protein design. *Prot. Sci.* **9:** 1391–1394.

Szilagyi, A. and Závodszky, P. 2000. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure Fold. Des.* **8:** 493–504.

Takai, K. and Sako, Y. 1999. A molecular view of archaeal diversity in marine and terrestrial hot water environments. *FEMS Microbiol. Ecol.* **28:** 177–188.

Tanner, J. J., Hecht, R. M., and Krause, K. L. 1996. Determinants of enzyme thermostability observed in the molecular structure of *Thermus aquaticus* D-glyceraldehyde-3–phosphate dehydrogenase at 2.5 Å resolution. *Biochemistry* **35:** 2597–2609.

Teter, S. A., Houry, W. A., Ang, D., Tradler, T., Rockabrand, D., Fischer, G., Blum, P., Georgopoulos, C., and Hartl, F. U. 1999. Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. *Cell* **97:** 755–765.

Thoma, R., Hennig, M., Sterner, R., and Kirschner, K. 2000. Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima*. *Structure Fold. Des.* **8:** 265–276.

Thomm, M., Madon, J., and Stetter, K. O. 1986. DNA-dependent RNA polymerases of the three orders of methanogens. *Biol. Chem. Hoppe-Seyler* **367:** 473–481.

Thompson, M. J. and Eisenberg, D. 1999. Transproteomic evidence of a loop-deletion mechanism for enhancing protein thermostability. *J. Mol. Biol.* **290:** 595–604.

Timasheff, S. N. 1993. The control of protein stability and association by weak interactions with water: How do solvents affect these processes? *Annu. Rev. Biophys. Biomol. Struct.* **22:** 67–97.

Timasheff, S. N. 1995. Solvent stabilization of protein structure. In: *Methods in Molecular Biology, Vol. 40: Protein Stability and Folding: Theory and Practice*, pp. 253–269. (Shirley, B. A., Ed.), Humana Press, Totowa.

Van de Vossenberg, J. L. C. M., Driessen, A. J. M., and Konings, W. N. 1998. The essence of being extremophilic: the role of the unique archaeal membrane lipids. *Extremophiles* **2:** 163–170.

Van den Burg, B., Vriend, G., Veltman, O. R., Venema, G., and Eijsink, V. G. 1998. Engineering an enzyme to resist boiling. *Proc. Natl. Acad. Sci. USA* **95:** 2056–2060.

Veinger, L., Diamant, S., Buchner, J., and Goloubinoff, P. 1998. The small heat-shock protein IbpB from *Escherichia coli* stabilizes stress-denatured proteins for subsequent refolding by a multichaperone network. *J. Biol. Chem.* **273:** 11032–11037.

Vetriani, C., Maeder, D. L., Tolliday, N., Yip, K. S., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H., and Robb, F. T. 1998. Protein thermostability above 100 °C: a key role for ionic interactions. *Proc. Natl. Acad. Sci. U S A* **95:** 12300–12305.

Viikari, L., Kantelinen, A., Sundquist, J., and Linko, M. 1994. Xylanases in biobleaching. From an idea to industy. *FEMS Microbiol. Lett.* **13:** 335–350.

Waldmann, T., Nimmesgern, E., Nitsch, M. Peters, J., Pfeifer, G., Müller, S., Kellermann, J., Engel, A., Hartl, F. U., and Baumeister, W. 1995. The thermosome of *Thermoplasma acidophilum* and its relationship to the eukaryotic chaperonin TRiC. *Eur. J. Biochem.* **227:** 848–856.

Wassenberg, D., Schurig, H., Liebl, W., and Jaenicke, R. 1997. Xylanase XynA from the hyperthermophilic bacterium *Thermotoga maritima*: Structure and stability of the recombinant enzyme and its isolated cellulose-binding domain. *Prot. Sci.* **6:** 1718–1726.

Wassenberg, D., Liebl, W., and Jaenicke, R. 2000. Maltose-binding protein from the hyperthermophilic bacterium *Thermotoga maritima*: stability and binding properties. *J. Mol. Biol.* **295:** 279–288.

Watanabe, Y., Motohashi, K., Taguchi, H., and Yoshida, M. 2000. Heat-inactivated proteins managed by DnaKJ-GrpE-ClpB

chaperones are released as a chaperonin-recognizable non-native form. *J. Biol. Chem.* **275:** 12388–12392.

Weber-Ban, E. U., Reid, B. G., Miranker, A. D., and Horwich, A. L. 1999. Global unfolding of a substrate protein by the Hsp100 chaperone ClpA. *Nature* **401:** 90–93.

Williams, J. C., Zeelen, J. P., Neubauer, G., Vriend, G., Backmann, J., Michels, P. A., Lambeir, A. M., and Wierenga, R. K. 1999. Structural and mutagenesis studies of leishmania triosephosphate isomerase: a point mutation can convert a mesophilic enzyme into a superstable enzyme without losing catalytic power. *Prot. Eng.* **12:** 243–250.

Wray, J. W., Baase, W. A., Lindstrom, J. D., Weaver, L. H., Poteete, A. R., and Matthews, B. W. 1999. Structural analysis of a non-contiguous second-site revertant in T4 lysozyme shows that increasing the rigidity of a protein can enhance its stability. *J. Mol. Biol.* **292:** 1111–1120.

Wright, H. T. 1991. Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins. *Crit. Rev. Biochem. Mol. Biol.* **26:** 1–52.

Xavier, K. B., da Costa, M. S., and Santos, H. 2000. Demonstration of a novel glycolytic pathway in the hyperthermophilic archaeon *Thermococcus zilligii* by ^{13}C-labeling experiments and nuclear magnetic resonance analysis. *J. Bacteriol.* **182:** 4632–4636.

Xiao, L. and Honig, B. 1999. Electrostatic contributions to the stability of hyperthermophilic proteins. *J. Mol. Biol.* **289:** 1435–1444.

Yamamoto, H., Hiraishi, A., Kato, K., Chiura, H. X., Maki, Y., and Shimizu, A. 1998. Phylogenetic evidence for the existence of novel thermophilic bacteria in hot spring sulfur-turf microbial mats in Japan. *Appl. Environ. Microbiol.* **64:** 1680–1687.

Yan, Z., Fujiwara, S., Kohda, K., Takagi, M., and Imanaka, T. 1997. *In vitro* stabilization and *in vivo* solubilization of foreign proteins by the beta subunit of a chaperonin from the hyperthermophilic archaeon *Pyrococcus* sp. strain KOD1. *Appl. Environ. Microbiol.* **63:** 785–789.

Yip, K. S., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., and Consalvi, V. 1995. The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure* **3:** 1147–1158.

Yoshida, T., Yohda, M., Iida, T., Maruyama, T., Taguchi, H., Yazaki, K., Ohta, T., Odaka, M., Endo, I., and Kagawa, Y. 1997. Structural and functional characterization of homo-oligomeric complexes of α and β chaperonin subunits from the hyperthermophilic archaeon *Thermococcus* strain KS-1. *J. Mol. Biol.* **273:** 635–645.

Závodszky, P., Kardos, J., Svingor, and Petsko, G. A. 1998. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl. Acad. Sci. USA* **95:** 7406–7411.

Zeikus, J. G., Vieille, C., and Savchenko, A. 1998. Thermozymes: biotechnoogy and structure-function relationships. *Extremophiles* **2:** 179–183.

Zhao, H., Giver, L., Shao, Z., Affholter, J. A., and Arnold, F. H. 1998. Molecular evolution by staggered extension process (StEP) *in vitro* recombination. *Nat. Biotechnol.* **16:** 258–261.

Ziemienowicz, A., Skowyra, D., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C., and Zylicz, M. Both the *Escherichia coli* chaperone systems, GroEL/GroES and Dnak/DnaJ/GrpE, can reactivate heat-treated RNA polymerase. Different mechanisms for the same acitivity.1993. *J. Biol. Chem.* **268:** 25425–25431.